

1258500

# THE UNITED STATES OF AMERICA

TO ALL TO WHOM THESE PRESENTS SHALL COME:

UNITED STATES DEPARTMENT OF COMMERCE

United States Patent and Trademark Office

*December 09, 2004*

THIS IS TO CERTIFY THAT ANNEXED HERETO IS A TRUE COPY FROM THE RECORDS OF THE UNITED STATES PATENT AND TRADEMARK OFFICE OF THOSE PAPERS OF THE BELOW IDENTIFIED PATENT APPLICATION THAT MET THE REQUIREMENTS TO BE GRANTED A FILING DATE.

APPLICATION NUMBER: 60/500,281

FILING DATE: *September 04, 2003*

RELATED PCT APPLICATION NUMBER: PCT/US04/27672

Certified by



Jon W Dudas

Acting Under Secretary of Commerce  
for Intellectual Property  
and Acting Director of the U.S.  
Patent and Trademark Office

BEST AVAILABLE COPY

PROVISIONAL APPLICATION

for

UNITED STATES LETTERS PATENT

on

METHODS OF IDENTIFYING AGENTS THAT INHIBIT THE GROWTH OF  
CANCER CELLS

by

Xiuyuan Hu, Richard Tritz, Henry Li, Ning Ke, Mirta  
Grifman, Cheryl Rogers, Kristin Defife, Cellia Habita,  
Wufang Fan, Kristina Rhoades, Philip Tan and Flossie Wong-  
Staal

Sheets of Drawings: 45

Docket No.: P-IMM 1008

CERTIFICATE OF MAILING BY "EXPRESS MAIL"

"EXPRESS MAIL" MAILING LABEL NUMBER: EF336774120US

DATE OF DEPOSIT: September 4, 2003

I HEREBY CERTIFY THAT THIS PAPER IS BEING  
DEPOSITED WITH THE UNITED STATES POSTAL SERVICE  
"EXPRESS MAIL POST OFFICE TO ADDRESSEE" SERVICE UNDER  
37 C.F.R. 1.10 ON THE DATE INDICATED ABOVE, AND IS  
ADDRESSED TO: COMMISSIONER FOR PATENTS, MAIL STOP: PROVISIONAL  
APPLICATION, P.O. BOX 1450, ALEXANDRIA, VA 22313-1450.

David I. Spolter  
(TYPED OR PRINTED NAME OR PERSON MAILING PAPER OR FEE)

David I. Spolter  
(SIGNATURE OF PERSON MAILING PAPER OR FEE)

Attorney

LAW OFFICE OF DAVID SPOLTER

1590 Coast Walk

La Jolla, California 92037

# **METHODS OF IDENTIFYING AGENTS THAT INHIBIT THE GROWTH OF CANCER CELLS**

## **FIELD OF THE INVENTION**

5

The present invention is directed to methods of identifying agents useful for inhibiting cancer cells.

The agents bind to one of the proteins described herein, or to a genes or mRNA encoding such protein. The

10 invention is also directed to methods for inhibiting cancer cells by administering an agent that binds to one of the proteins described herein, or to a gene or mRNA encoding such protein.

15

## **BACKGROUND OF THE INVENTION**

Cancer is a major cause of mortality worldwide.

Despite advancements in diagnosis and treatment, there

20 remains a great need for novel methods of treating cancer and for identifying novel agents that inhibit cancer cells. The present invention satisfies this need and provides additional benefits as well.

25

## **SUMMARY OF THE INVENTION**

The present invention centers on the discovery that knockdown of certain proteins expressed in cancer

cells, as well as the genes or mRNA encoding such

30 proteins, results in the death of the cancer cells, reduction in their size, a decrease in their growth or an increase in sensitization to undergo apoptosis.

Accordingly, the present invention provides methods for identifying agents useful in inhibiting cancer cells,

such methods based upon the binding of such agents a protein of the invention, or to a gene or mRNA encoding such protein. Such methods include introducing an agent into cancer cells in which one or more of the proteins of the invention is expressed, and determining the effect of such agent on the cells. An agent that effectively binds to one of the proteins of the invention and thereby causes a decrease in cancer cell proliferation, or an increase in cell death (apoptosis), or otherwise decreases cancerous growth, will be useful for the treatment of cancer. Representative agents include antisense oligonucleotides, ribozymes, siRNAs, monoclonal and polyclonal antibodies, and small organic molecules.

#### **BRIEF DESCRIPTION OF THE DRAWINGS**

Figure 1 shows the general structure and nucleotide sequence of a hairpin ribozyme (large case lettering) (SEQ ID NO:13) and its interaction with a substrate RNA (small case lettering) (SEQ ID NO:14).

Figure 2 shows the general structure and nucleotide sequence of a hammerhead ribozyme (large case lettering) (SEQ ID NO:15) and its interaction with a substrate RNA (small case lettering) (SEQ ID NO:16).

Figure 3 shows the DNA sequence of stearoyl-CoA desaturase ("SCD;" SEQ ID NO:1).

Figure 4 shows the DNA sequence of carbonic anhydrase XII ("CA12;" SEQ ID NO:2).

Figure 5 shows the DNA sequence of phosphoinositide-3-kinase, regulatory subunit 4, p150 ("PIK3R4;" SEQ ID NO:3).

5           Figure 6 shows the DNA sequence of phospholipase D3 ("PLD3;" SEQ ID NO:4).

Figure 7 shows the DNA sequence of heat shock protein 60 ("HSPD1;" SEQ ID NO:5).

10

Figure 8 shows the DNA sequence of zipper protein kinase ("ZPK;" SEQ ID NO:6).

Figure 9 shows the amino acid sequence of SCD  
15 (SEQ ID NO:7).

Figure 10 shows the amino acid sequence of CA12 (SEQ ID NO:8).

20           Figure 11 shows the amino acid sequence of PIK3R4 (SEQ ID NO:9).

Figure 12 shows the amino acid sequence of PLD3 (SEQ ID NO:10).

25

Figure 13 shows the amino acid sequence of HSPD1 (SEQ ID NO:11).

Figure 14 shows the amino acid sequence of ZPK  
30 (SEQ ID NO:12).

Figure 15 shows phenotypic data (increase in Fas-induced apoptosis) obtained from DLD-1 cells when

transfected with siRNA constructs targeted against SCD  
(SEQ ID NO:1).

Figure 16 shows SCD mRNA knockdown data  
5 obtained from DLD-1 cells when transfected with siRNA  
constructs targeted against SCD (SEQ ID NO:1).

Figure 17 shows phenotypic data (increase in  
Fas-induced apoptosis) obtained from A2058 cells when  
10 transfected with siRNA constructs targeted against SCD  
(SEQ ID NO:1).

Figure 18 shows SCD mRNA knockdown data  
obtained from A2058 cells when transfected with siRNA  
15 constructs targeted against SCD (SEQ ID NO:1).

Figure 19 shows phenotypic data (increase in  
TRAIL-induced apoptosis) obtained from A2058 cells  
transfected with an siRNA (SEQ ID NO:17) targeting SCD  
20 (SEQ ID NO:1), as measured by a DNA fragmentation ELISA  
assay.

Figure 20 shows phenotypic data (increase in  
Fas- and TRAIL-induced apoptosis) obtained from A2058  
25 cells transfected with an siRNA (SEQ ID NO:17) targeting  
SCD (SEQ ID NO:1); the data for TRAIL-induced apoptosis  
is based upon Caspase 3/7 activity.

Figure 21 shows the data from three different  
30 transfection experiments using SEQ ID NOS:17 to 19, siRNA  
constructs targeted against SCD (SEQ ID NO:1), in normal  
(i.e., non-cancerous) cells.

Figure 22 shows virtual northern data for SCD (SEQ ID NO:1) in normal and breast cancer cell lines.

Figure 23 shows phenotypic data (increase in Fas-induced apoptosis) obtained from DLD-1 cells and summarizes the results from three independent transfection experiments using siRNA constructs targeted against CA12 (SEQ ID NO:2).

Figure 24 shows CA12 mRNA knockdown data obtained from DLD-1 cells using siRNA constructs targeted against CA12 (SEQ ID NO:2).

Figure 25 shows phenotypic data (increase in Fas-induced apoptosis) obtained from AsPC-1 cells using siRNA constructs targeted against CA12 (SEQ ID NO:2).

Figure 26 shows CA12 mRNA knockdown data obtained from AsPC-1 cells using siRNA constructs targeted against CA12 (SEQ ID NO:2).

Figure 27 shows phenotypic data (reduced survival) of A172 cells after transfection with siRNAs targeting PI3KR4 (SEQ ID NO:3).

Figure 28 shows the expression of PIK3R4 (SEQ ID NO:3) in glioblastoma, colon, and brain normal and cancer cell lines.

Figure 29A shows phenotypic data (reduced survival) of A172 cells after transfection with siRNAs targeting PLD3 (SEQ ID NO:4); Figure 29B shows knockdown of PLD3 mRNA after transfection with siRNAs targeting PLD3.

Figure 30 shows the expression of PLD3 (SEQ ID NO:4) in normal and cancer cell lines.

5           Figure 31 shows the expression levels of HSPD1 (SEQ ID NO:5) in several experimental cell lines: Hela (cervical cancer); AsPC-1 (pancreatic cancer); DLD-1 (colon cancer), PC3 (prostate cancer); T47D (breast cancer), A2058 (melanoma) and U87 (glioma).

10

Figure 32 shows phenotypic data (reduced cell viability) of A2058 cells after transfection with siRNAs targeting HSPD1 (SEQ ID NO:5); Figure 32A shows cells treated with r 1% serum; Figure 32B shows cells treated  
15 with 1% serum + brefeldin A (0.05µg/ml).

Figure 33 shows phenotypic data (reduced cell viability) of AsPC-1 cells after transfection with siRNAs targeting HSPD1 (SEQ ID NO:5); Figure 33A shows cells  
20 treated with 1% serum; Figure 33B shows cells treated with 1% serum + brefeldin A (0.05µg/ml).

Figure 34 shows phenotypic data (reduced cell viability) of PC3 cells after transfection with siRNAs  
25 targeting HSPD1.

Figure 35 shows phenotypic data (increase in apoptosis as measured by Caspase 3/7 activity) of A2058 cells after transfection with siRNAs targeting HSPD1.

30

Figure 37 shows the alignment of two isoforms (SEQ ID NOS:6 and 40, variants 2 and 1, respectively) of ZPK.



Figure 38 shows ZPK (SEQ ID NO:6) mRNA levels in several cancer cell lines.

5           Figure 39 shows siRNA mediated ZPK (SEQ ID NO:6) mRNA downregulation and cancer growth reduction in Hela cells. Figure 39A shows anchorage dependent growth reduction. Figure 39B shows anchorage independent growth reduction. Figure 39C shows reduction of ZPK mRNA based  
10 upon real-time RT-PCR (Taqman®) analysis.

Figure 40 shows ZPK (SEQ ID NO:6) siRNA-mediated reduction in anchorage independent growth of DLD-1 (Figure 40A), AsPC1 (Figure 40B) and PC3 (Figure  
15 40C) cells.

Figure 41 shows the DNA sequence of variant 1 of ZPK (SEQ ID NO:40).

20           Figure 42 shows the amino acid sequence of variant 1 of ZPK (SEQ ID NO:41).

#### **DETAILED DESCRIPTION OF INVENTION**

25           The present invention centers on the discovery of the correlation between the knockdown of various proteins, as well as the genes and mRNA encoding such proteins, and the treatment of cancer.

30           Specifically, the invention provides a method of identifying a molecule useful for the treatment of cancer by introducing the molecule into cancer cells, where it binds to a protein encoded by one of the genes

of the invention, SCD, CA12, PIK3R4, PLD3, HSPD1 or ZPK (SEQ ID NOS:1, 2, 3, 4, 5 or 6, respectively) or to the mRNA of SEQ ID NOS:1, 2, 3, 4, 5 or 6. The method further includes measuring the level of "inhibition" of the cancer cells. "Inhibition" means any one or more of the following: a) a decrease in cancer cell division (proliferation); b) an increase in the sensitivity of the cells to undergo apoptosis; or c) an increase in cell death (apoptosis).

Examples of cancer cells useful for the practice of the invention include HeLa cells, A2058 cells, DLD-1 cells, T47D cells, ASPC-1 cells and JH-MG cells.

The invention further provides a method of identifying a molecule that inhibits cancer cells by introducing the molecule into cells, where the molecule binds to a domain of a protein of the invention, specifically a domain of SEQ ID NOS:7, 8, 9, 10 or 12. The domains of the invention include residues 96 to 321 of SEQ ID NO:7, which is the fatty acid desaturase domain of the SCD protein; residues 32 to 289 of SEQ ID NO:8, which is the carbonate anhydrase domain of the CA12 protein; residues 26 to 320 of SEQ ID NO:9, which is the protein kinase domain of the PIK3R4 protein; residues 143 to 170 of SEQ ID NO:10, which is an active site motif of the PLD protein; residues 358 to 384 of SEQ ID NO:10, which is an active site motif of the PLD protein; residues 47 to 550 of SEQ ID NO:11, which is the chaperone portion of the HSPD1 protein; residues 158 to 405 of SEQ ID NO:12, which is the protein kinase domain of variant 2 of the ZPK protein; and residues 175 to 414 of SEQ ID NO:41, which is the protein kinase domain of

variant 1 of the ZPK protein. The method of the invention further provides measuring the level of inhibition of the cells.

5           Measurement of cancer cell inhibition may be done, for example, by means of an apoptosis assay, where an increase in the level of apoptosis indicates that the molecule introduced into the cell system inhibits the cancer cells. Measurement of inhibition may also be done  
10 by means of an assay that measures cell proliferation, where a decrease in the rate of cell division indicates that the molecule inhibits the cancer cells. In addition, measurement of cancer cell inhibition may also be done by assessing reduction in the growth of cancer cells in  
15 liquid media (anchorage dependent growth) and/or in soft agar (anchorage independent growth), where a measurable decrease in the rate of growth indicates that the molecule inhibits the cancer cells.

20           The invention further provides a method of inhibiting cancer cells by introducing into the cells a molecule that binds to a compound comprising SEQ ID NOS:1, 2, 3, 4, 5 or 6, or comprising the RNA correlate of SEQ ID NOS:1, 2, 3, 4, 5 or 6.

25

          As shown in Example 1, exemplary of such a molecule is an siRNA. Introduction into DLD-1 and A2058 melanoma cells of siRNA molecules (SEQ ID NOS:17 and 18) targeting SCD (SEQ ID NO:1) caused reduction of the mRNA  
30 of SCD. See Figures 15 and 17. Moreover, these siRNA constructs were shown to cause an increase of apoptosis in these cancer cell lines, even when such cancer cell lines are normally resistant to apoptosis. See Figures 16 and 18 to 20. In addition, siRNA constructs (SEQ ID

NOS:17 to 19) targeted against SCD (SEQ ID NO:1) did not activate apoptosis in normal (i.e., non-cancerous) cells. See Figure 21. Thus, these results show that these constructs selectively activate apoptosis only in cancer  
5 cells.

Moreover, the introduction into cancer cell lines of two siRNA molecules (SEQ ID NOS:22 and 24) targeting CA12 (SEQ ID NO:2) each caused reduction of the  
10 mRNA of CA12 and increased apoptosis in the cancer cell lines. Specifically, CA12-3 and CA12-5 (SEQ ID NOS:22 and 24) were shown to do so in DLD-1 cells (see Example 2 and Figures 23 and 24) and AsPC-1 cells (see Example 2 and Figures 25 and 26).

15

Furthermore, four other siRNA constructs of the invention (SEQ ID NOS:25 to 28) against another gene of the invention, PIK3R4 (SEQ ID NO:3), increased apoptosis in the A172 glioblastoma cell line. See Example 3 and  
20 Figure 27.

In addition, four separate siRNA constructs (SEQ ID NOS:29 to 32) targeting another gene of the invention, PLD3 (SEQ ID NO:4), increased the level of  
25 sensitivity to apoptosis in the glioblastoma cell line A172 cells (see Figure 29A) and knocked down the level of PLD3 mRNA (see Figure 29B).

Moreover, four separate siRNA constructs (SEQ  
30 ID NOS:33 to 36) were made against the mRNA of a gene of the invention, HSPD1 (SEQ ID NO:5). Three of these constructs (SEQ ID NOS:33 to 35) decreased cell viability in A2058 cells (see Example 5 and Figure 32A), and all four constructs increased sensitivity of these cells to

apoptosis (Figure 32B). Further, two of these constructs (SEQ ID NOS:33 and 34) reduced the viability and increased sensitivity to apoptosis in both AsPC1 cells (Figures 33A and 33B) and PC3 prostate cancer cells (Figure 34), and increased the sensitivity to apoptosis in A2058 cells (Figure 35).

In yet another example, two siRNA constructs of the invention (SEQ ID NOS:37 and 39) made to target a gene of the invention, ZPK (SEQ ID NO:6), caused significant growth reduction in anchorage dependent medium (Figure 39A), and even greater growth reduction in anchorage independent medium (Figure 39B). A correlative knockdown of ZPK mRNA was also shown (Figure 39C). See Example 6. Similar results regarding three siRNA constructs (SEQ ID NOS:37 to 39) were achieved in DLD-1, AsPC-1 and PC3 cell lines (Figures 40A, 40B and 40C).

In addition to siRNA constructs, a molecule of the invention also includes ribozymes, antisense molecules, antagonizing antibodies and small organic molecules that target or bind to one of the genes of the invention.

An example of an apoptosis assay is the Annexin-V binding assay. This assay is based on the relocation of phosphatidylserine to the outer cell membrane. Viable cells maintain an asymmetric distribution of different phospholipids between the inner and outer leaflets of the plasma membrane. Choline-containing phospholipids such as phosphatidylcholine and sphingomyelin are primarily located on the outer leaflet of viable cells and aminophospholipids such as phosphatidylethanolamine and phosphatidylserine (PS) are

found at the cytoplasmic (inner) face of viable cells. The distribution of phospholipids in the plasma membrane changes during apoptosis. In particular, PS relocates from the cytoplasmic face to the outer leaflet so called  
5 PS exposure. The extent of PS exposure can distinguish apoptotic cells from the non-apoptotic cells.

Annexin-V is a 35-36 kDa calcium-dependent phospholipid binding protein with high affinity for PS  
10 (kDa ~  $5 \times 10^{-10}$  M). When labeled with a fluorescent dye, Annexin-V can be used as a sensitive probe for PS exposure on the outer leaflet of the cell membrane. The binding of Annexin-V conjugates such as Annexin-V FITC to cells permits differentiation of apoptotic cells  
15 (Annexin-V positive) from non-apoptotic cells (Annexin-V negative). Annexin-V binding is observed under two conditions. The first condition is observed in cells midway through the apoptosis pathway. Phosphatidylserine translocates to the outer leaflet of the cell membrane.  
20 The second condition is observed in very late apoptosis or when the cells become necrotic and membrane permeabilization occurs. This membrane permeabilization allows Annexin-V to enter cells and bind to phosphatidylserine on the cytoplasmic face of the  
25 membrane. Since other causes besides apoptosis can result in necrosis, it is important to distinguish between necrotic and apoptotic cells. Membrane permeabilization also permits entry of other materials to the interior of the cell, including  
30 the fluorescent DNA-binding dye propidium iodide. Utilizing dual staining methodology, apoptotic populations can be distinguished from necrotic populations. For example, using the Annexin V-propidium iodide (PI) double staining regime, three populations of

cells are distinguishable in twocolor flow cytometry.  
See Boersma, et al., *Cytometry*, 24:123-130 (1996);  
Martin, et al., *J. Exp. Med.*, 182:1545-1556 (1995).

5           Another example of an apoptosis assay is the  
caspase 3/7 assay. Briefly, caspases are synthesized as  
inactive pro-enzymes or pro-caspases. In apoptosis, the  
pro-caspases are processed by proteolytic cleavage to  
form active enzymes. For example, caspase-3 exists in  
10 cells as an inactive 32 kDa proenzyme, called pro-  
caspase-3. Pro-caspase-3 is cleaved into active 17 and  
12 kDa subunits by upstream proteases to become active  
caspase-3. Caspases-2, -8, -9 and -10 are classified as  
signaling or "upstream" in the apoptosis pathway because  
15 long prodomains allow association with cell surface  
receptors such as FAS (CD95), TNFR-1 (CD120a), DR-3 or  
CARD domains. This observation suggests a proteolytic  
cascade as a mechanism for signaling. A proteolytic  
cascade exists that would activate the terminal event  
20 required for apoptosis in a way similar to that of the  
coagulation cascade seen with the closely related family  
of serine proteases. For example, caspase-4 activates  
pro-caspase-1; caspase-9 activates pro-caspase-3; and  
caspase-3 cleaves pro-caspase-6 and pro-caspase-7.  
25 Caspases play a critical role in the execution phase of  
apoptosis. Important targets of caspases include  
cytoplasmic and nuclear proteins such as keratin 18, poly  
ADP ribose polymerase (PARP) and lamins. Overexpression  
of caspase-3 induces apoptosis. Through the use of  
30 synthetic peptides, caspases have been divided into three  
groups based on the four amino acids amino-terminal to  
their cleavage site. Caspases-1, -4 and -5 prefer  
substrates containing the sequence WEXD (where X is  
variable). Caspases-2, -3 and -7 prefer the sequence

DEXD. Caspases 6, 8 and 9 are the least demanding but have demonstrated a preference for cleaving of substrates containing either LEXD or VEXD. Because these sequences correspond to known cleavage sites of caspase targets, systems to study caspase cleavage activity have been developed. The measurement of caspase enzyme activity with fluorometric and colorimetric peptide substrates and the detection of caspase cleavage using antibodies to caspases allows the study of the apoptosis processes or screening of therapeutic agents which promote or prevent apoptosis. A typical assay would involve the cleavage of a fluorescent substrate peptide to quantitate activity. The substrate, DEVD-AFC, is composed of the fluorophore, AFC (7-amino-4-trifluoromethyl coumarin), and a synthetic tetrapeptide, DEVD (Asp-Glu-Val-Asp), which is the upstream amino acid sequence of the Caspase-3 cleavage site in PARP. DEVD-AFC emits blue light ( $\lambda_{\text{max}} = 400$  nm). Upon cleavage of the substrate by Caspase-3 or related caspases, with the excitation wavelength set to 400 nm, free AFC emits a yellow-green fluorescence ( $\lambda_{\text{max}} = 505$  nm) which can be quantified using a spectrofluorometer or a fluorescence microtiter plate reader. Comparison of the fluorescence of AFC from apoptotic samples with an uninduced control allows determination of the increase in caspase-3 activity. See Jaeschke, et al. J. Immunol., 160:3480-3486 (1998); Talanian, et al., J. Biol. Chem., 272:9677-9682 (1997).

Yet another example of an apoptosis assay is the TUNEL assay. Briefly, cell death by apoptosis is characterized by DNA fragmentation in 200-250 and/or 30-50 kilobases. Further internucleosomal DNA fragmentation in 180-200 base pairs may also occur. Such characteristics have been used to distinguish apoptotic



cells from normal or necrotic cells. To detect apoptotic cells, whatever the pattern of DNA fragmentation, the TUNEL (Terminal deoxynucleotidyl transferase (TdT) mediated dUTP Nick End Labeling) method is commonly  
5 utilized. One of the most easily measured features of apoptotic cells is the break-up of the genomic DNA by cellular nucleases. These DNA fragments can be extracted from apoptotic cells and result in the appearance of "DNA  
10 laddering" when the DNA is analyzed by agarose gel electrophoresis. The DNA of non-apoptotic cells, which remains largely intact, does not display this "laddering" on agarose gels during electrophoresis. The large number of DNA fragments appearing in apoptotic cells results in a multitude of 3'-hydroxyl ends in the DNA. This  
15 property can be used to identify apoptotic cells by labeling the 3'-hydroxyl ends with bromolabeled deoxyuridine triphosphate nucleotides (Br-dUTP). The enzyme terminal deoxynucleotidyl transferase (TdT) catalyzes a template independent addition of  
20 deoxyribonucleoside triphosphates to the 3'-hydroxyl ends of double- or single-stranded DNA with either blunt, recessed or overhanging ends. See Li and Darzynkiewicz, Cell Prolif., 28:572-579 (1995).

25 In another apoptosis assay, the cell death ELISA detects the same endpoint as the TUNEL assay, DNA fragmentation. However, in the cell death ELISA assay, the histone complexed DNA fragments are measured directly by antibodies in an ELISA assay. See Piro, et al.,  
30 Metabolism, 51:1340-7 (2002); Facchiano et al., Exp. Cell Res., 271:118-29 (2001); Horigome et al., Immunopharmacology, 37:87-94 (1997).

The invention further provides a method of decreasing cell proliferation, comprising introducing into said cell an agent that down-modulates the activity of SEQ ID NO: 1, 2, 3, 4, 5 or 6. The agent can be a  
5 ribozyme, an antisense molecule, an antagonizing antibody or an siRNA construct, as discussed above.

An example of a cell proliferation assay is the alamarBlue™ assay. Other examples include the Natural  
10 Red, methylene blue and tetrazolium/formazan assays. See also biosource.com; Alley et al., Cancer Res., 48:589-601 (1988); Elliot and Auersperg, Biotech. Histochem., 68:29-35 (1993); and Scudiero et al., Cancer Res., 48:4827-33 (1988). Another example of a cell proliferation assay is  
15 based on the cleavage of the terazolium salt WST-1(Roche Diagnostics, GmbH; available at the Roche website, [biochem.roche.com/pack-insert/1644807.pdf](http://biochem.roche.com/pack-insert/1644807.pdf)).

Another example of an assay that measures a  
20 difference in cell proliferation (or tumorigenicity) is a soft agar assay, as described in Example 6. See also Bergo et al., Mol. and Cell. Biol., 22:171-181 (2002); Zeng et al., Cancer Res., 62:3538:3543 (2002).

25 The invention further provides a method of identifying a molecule that inhibits cancer cells by introducing the molecule into cells, where the molecule down-modulates the RNA correlate of SEQ ID NOS:1, 2, 3, 4, 5 or 6. The method further provides measuring the  
30 level of down-modulation of the compound, where an increase in level of down-modulation indicates that said molecule inhibits cancer cells.

Measurement of down-modulation can be made by a reporter assay using a reporter gene operably linked to a nucleic acid encoding any of the proteins of the invention. Reporter genes can express proteins such as  $\beta$ -lactamase, luciferase, green fluorescent protein,  $\beta$ -galactosidase, secreted alkaline phosphatase, human growth hormone and chlororamphenicol acetyltransferase.

The invention also provides a method of identifying a compound that inhibits cancer cells by introducing the compound into cells, where it down-modulates SEQ ID NOS:7, 8, 9, 10, 11 or 12. The method further provides measuring the level of down-modulation of the compound, where an increase in level of down-modulation indicates that said molecule inhibits cancer cells. Down-modulation can be measured, for example, by an immunoassay using an antibody specific to said compound. Such an immunoassay can be, for example, an immunofluorescence, immunochemistry or immunoprecipitation assay.

The isolated nucleic acid molecules of the invention include nucleotide sequences encoding ribozymes. The ribozymes of the invention have "substrate binding sequences" that hybridize to and cleave complementary sequences of the mRNA of the genes disclosed herein and, therefore, increase inhibition of cancer cells. Included within the scope of the invention are expression vectors encoding the ribozymes, cells containing the vectors and cells expressing the ribozymes. In addition, a ribozyme of the present invention can be introduced directly into a cell, i.e., without the use of a vector.

The ribozymes of the invention are catalytic RNA molecules that bind to the target nucleic acid molecules and cleave them, thereby impairing their ability to function in cancer growth and development.

5 The ribozymes of the invention are identified and selected by methods described herein. They may be "hairpin" ribozymes, "hammerhead" ribozymes or any other type of ribozyme known in the art.

10 Figure 1 illustrates the basic structure and nucleotide sequence (shaded, in uppercase letters) of a hairpin ribozyme (SEQ ID NO:13) and its relationship to the complementary nucleotide sequence (lowercase letters; SEQ ID NO:14) of a target substrate (N or n = any  
15 nucleotide). A hairpin ribozyme consists of a 50 to 54 nucleotide RNA molecule, with the non-substrate binding sequence beginning from the 5' end at nucleotide position 17. It folds into a 2-dimensional structure that resembles a hairpin, consisting of two helical domains  
20 (Helix 3 and 4) and 3 loops (Loop 2, 3 and 4) and two additional helices (Helix 1 and 2), which form between the ribozyme and the substrate. The length of Helix 2 is fixed at 4 base pairs and the length of Helix 1 typically varies from 6 to 10 base pairs. Recognition of the  
25 substrate nucleotides by the ribozyme occurs via Watson-Crick base pairing, with typical substrate recognition sites having the structure 5'-GUC-3' or 5'-GUA-3'. For maximal activity, the RNA target substrate will contain a GUC in a loop that is opposite Loop 1 with  
30 cleavage occurring immediately 5' of the G as indicated by an arrow. The catalytic, but not substrate binding, activity of a hairpin ribozyme can be disabled by mutating the 5'-AAA-3' in Loop 2 to 5'-CGU-3'.

The general structure of a hammerhead ribozyme is shown in Figure 2 (SEQ ID NO:15) along with its target RNA sequence (SEQ ID NO:16). Hammerhead ribozymes suitable for use within the present invention preferably  
5 recognize the sequence NUH, wherein N is any of G, U, C, or A and H is C, U, or A. It will be appreciated that the recognition sites of the hairpin ribozyme of the present invention (5'-GUC-3' or 5'-GUA-3') are a subset of the hammerhead recognition sites (5'-NUH-3'), such  
10 that all hairpin recognition sites are by definition also hammerhead recognition sites, although the converse is not true.

The invention also includes a chimeric ribozyme  
15 based on the ribozyme's substrate binding sequence. A chimeric hammerhead ribozyme (i.e., a RNA/DNA hybrid) is designed to recognize the appropriate NUH sequence for cleavage. Generally, most or all of the binding arms and stem loop comprise DNA. By contrast, generally, the  
20 catalytic domain (shown in Figure 2 between the binding arms and stem loop) comprises RNA.

Modification of the base composition at the stem loop or catalytic domain regions can increase the  
25 catalytic activity of the ribozyme, as assayed by in vitro cleavage. See WO 00/32765. Modification at the 2-position of the sugar of the base, for example, substituting -OCH<sub>3</sub> at this position of an RNA base, can increase the stability of the ribozyme. Other  
30 stabilizing substitutions include -OC<sub>1-6</sub>Alkyl, -F or other halogens, amino, azido, nitro and phenyl. See U.S. Pat. No. 5,298,612.

Similarly, ribozyme activity can be optimized by altering the length of the ribozyme binding arms, or chemically synthesizing ribozymes with modifications that prevent their degradation by serum ribonucleases (see  
5 e.g., Int. Pat. Appl. Publ. No. WO 92/07065; Perrault et al., 1990; Pieken et al., 1991; Usman and Cedergren, 1992; Int. Pat. Appl. Publ. No. WO 93/15187; Int. Pat. Appl. Publ. No. WO 91/03162; Eur. Pat. Appl. Publ. No. 92110298.4; U.S. Pat. No. 5,334,711; and Int. Pat. Appl.  
10 Publ. No. WO 94/13688, which describe various chemical modifications that can be made to the sugar moieties of enzymatic RNA molecules), modifications which enhance their efficacy in cells, and removal of stem II bases to shorten RNA synthesis times and reduce chemical  
15 requirements.

The term, "substrate binding sequence" of a ribozyme, as used herein, refers to that portion of the ribozyme which base pairs with a complementary sequence  
20 (referred to herein as a "ribozyme sequence tag" or "RST") of a target nucleic acid. For example, the 16 nucleotides at the 5' end of the sequence of Figure 1 (SEQ ID NO:13) represent the general formula for a hairpin ribozyme substrate binding sequence. This  
25 includes the two arms that form helixes with the target and any necessary nucleotides between these two arms that may be required for the ribozyme function (e.g., AAGA or AAGC). The general formula for the substrate binding sequence of a hammerhead ribozyme is shown in Figure 2  
30 (SEQ ID NO:15) with the substrate binding sequence shown aligned with the complementary sequence (RST) of the target RNA (SEQ ID NO: 16).

Because of the basic similarity in the structure of all hairpin ribozymes, they can be modified to obtain a ribozyme having a specific substrate binding sequences of choice. For example, the substrate binding  
5 sequence of a "GUC ribozyme", which cleaves an RNA having the sequence 5'-NNNNN\*GUCNNNNNNNN (SEQ ID NO:45) may be modified to that of a "GUA ribozyme," which cleaves an RNA having the sequence 5' - NNNNN\*GUANNNNNNNN (SEQ ID NO:46), by changing the base at position 9 from the 5'  
10 end of the substrate binding sequence of the ribozyme. In both of these sequences, N is any of G, U, C, or A; and the asterisk indicates the site where cleavage of the target RNA occurs.

15 A preferred "GUC hairpin ribozyme" has a substrate binding sequence with the general formula 5'-(N)<sub>(6-10)</sub>AGAA(N)<sub>4</sub>-3' (SEQ ID NO:47), where N can be either G, T, C, or A. A preferred "GUA hairpin ribozyme" has a substrate binding sequence with the general formula 5'-  
20 (N)<sub>(6-10)</sub>CGAA(N)<sub>4</sub>-3' (SEQ ID NO:48), where N can be either G, T, C, or A. In the case of the specific ribozyme substrate binding sequences disclosed herein based on these formulas, the sequences also include variations where no more than two nucleotides differ at any of  
25 positions 1-5 from the 5' end of the sequence. This means that one may change one or two bases within the first 5 nucleotides at the 5' end of a substrate binding sequence and still retain the functional activity of the ribozyme.

30

Similarly, once the substrate binding sequence of a specific hairpin ribozyme has been identified, it is relatively easy to engineer an equivalent substrate binding sequence for a hammerhead ribozyme. The general

structure of the substrate binding sequence of a hammerhead includes three or four to nine bases at the 5' end of the ribozyme's binding arm and three or four to nine bases at the 3' end of the other binding arm. The following approach, for example, may be used: 1) identify the ribozyme sequence tag (RST) of the RNA target of the specific hairpin ribozyme of interest; 2) specify the first six to nine nucleotides at the 5' end of the hammerhead as complementary to the first six to nine nucleotides at the 3' end of the RST; 3) specify the first 5 nucleotides at the 3' end of the hammerhead as complementary to nucleotides at the 5' end of the RST; and 4) specify nucleotides at positions 6 and 7 from the 3' end of the hammerhead as complementary to the RST, while base 8 from the 3' end is an A.

It should therefore be understood that a ribozyme of the present invention can comprise a) 3, 4, 5 or 6 contiguous bases of the substrate binding sequence as part of one binding arm of the ribozyme; and b) 3, 4, 5 or 6 contiguous bases of any of the remaining contiguous bases of that ribozyme substrate binding sequence as part of the other binding arm of the ribozyme.

As used herein, the term "unmodified base" means one of the bases adenine, guanine, cytosine, uracil or thymine attached to the 1-carbon of the sugar (deoxyribose or ribo-furanose), with a phosphate bound to the 5-carbon of the sugar. Bases are bound to each other via phosphodiester bonds between the 3-carbon of one base and the 5-carbon of the next base.



As used herein, the term "modified base" means any base whose chemical structure is modified as follows. Adenine can be modified to result in 6-dimethyl-amino-purine, 6-methyl-amino-purine, 2-amino-purine, 2,6-  
5 diamino-purine, 6-amino-8-bromo-purine or 6-amino-8-fluoro-purine. Cytosine can be modified to result in 5-bromo-cytosine, 5-fluoro-cytosine, N,N-dimethyl-cytosine, N-methyl-cytosine, 2-thio-cytosine or 2-pyridone. Guanine can be modified to result in 8-bromo-guanine, 8-  
10 fluoroguanine, 2-amino-purine, hypoxanthine (inosine), 7-deaza-guanine or 6-thio-guanine. Uracil can be modified to result in 3-methyl-uracil, 5,6-dihydro-uracil, 4-thio-uracil, thymine, 5-bromo-uracil, 5-iodo-uracil or 5-fluoro-uracil. Thymine can be modified to result in 3-  
15 methyl-thymine, 5,6-dihydro-thymine, 4-thio-thymine, uracil, 5-bromo-uracil, 5-iodo-uracil or 5-fluoro-uracil. Methods of making such modifications as well as other modifications, such as halogen, hydroxy, amine, alkyl, azido, nitro and phenyl substitutions are disclosed in  
20 U.S. Pat. No. 5,891,684; and U.S. Pat. No. 5,298,612. The present invention encompasses sequences where one or more bases are modified.

In addition, the sugar moiety of a base can be  
25 modified as disclosed above regarding bases of a hammerhead ribozyme. The present invention encompasses sequences where one or more bases are so modified.

As used herein, the term "nucleic acid" or  
30 "nucleic acid molecule" refers to deoxyribonucleotides or ribonucleotides, oligomers and polymers thereof, in either single- or double-stranded form. Unless specifically limited, the term encompasses nucleic acids containing known analogues of natural nucleotides that

have similar binding properties as the reference nucleic acid. For example, as disclosed herein, such analogues include those with substitutions, such as methoxy, at the 2-position of the sugar moiety. Unless otherwise  
5 indicated by the context, the term is used interchangeably with gene, cDNA and mRNA encoded by a gene.

As used herein, the phrase "a nucleotide  
10 sequence encoding" refers to a nucleic acid which contains sequence information, for example, for a ribozyme, mRNA, siRNA, and the like, or for the primary amino acid sequence of a specific protein or peptide. In reference to a ribozyme, unless otherwise indicated, the  
15 explicitly specified encoding nucleotide sequence also implicitly covers sequences that do not materially effect the specificity of the ribozyme for its target nucleic acid. In reference to a protein or peptide, unless otherwise indicated, the explicitly specified encoding  
20 nucleotide sequence also implicitly encompasses variations in the base sequence encoding the same amino acid sequence (e.g., degenerate codon substitutions). The invention also contemplates proteins or peptides with conservative amino acid substitutions. The identity of  
25 amino acids that may be conservatively substituted is well known to those of skill in the art. Degenerate codons of the native sequence or sequences may be chosen to conform with codon preference in a specific host cell.

30 As used herein, the term "RNA correlate" of a given DNA sequence means that sequence with "U" substituted for "T," with the entire sequence in ribonucleic acid form. The present invention encompasses the RNA correlates of SEQ ID NOS:1, 2, 3, 4, 5 and 6.

The terms "sequence similarity", "sequence identity", or "percent identity," in the context of two or more nucleic acids or polypeptide sequences, refer to two or more sequences or subsequences that are, when optimally aligned with appropriate nucleotide insertions or deletions, the same or have a specified percentage of amino acid residues or nucleotides that are the same (i.e., 50% identity, 65%, 70%, 75%, 80%, preferably 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or higher identity to an amino acid sequences such as SEQ ID NOS:7, 8, 9, 10, 11 or 12 (or domains thereof), or a nucleotide sequence such as SEQ ID NOS:1, 2, 3, 4, 5 or 6 (or RNA correlates thereof), when compared and aligned for maximum correspondence over a comparison window, or designated region as measured using one of the following sequence comparison algorithms or by manual alignment and visual inspection. This definition also refers to the complement of a test sequence. Preferably, the identity exists over a region that is at least about 25 amino acids or nucleotides in length, or more preferably over a region that is 50-100 amino acids or nucleotides in length. These relationships hold, notwithstanding evolutionary origin (Reeck et al., *Cell*, 50:667 (1987)). When the sequence identity of a pair of polynucleotides or polypeptides is greater or equal to 65%, the sequences are said to be "substantially identical."

Alternatively, substantial identity will exist when a nucleic acid will hybridize under selective hybridization conditions, to a strand or its complement. Typically, selective hybridization will occur when there is at least about 55% homology over a stretch of at least about 14 nucleotides, more typically at least about 65%,

preferably at least about 75%, and more preferably at least about 90%. See, Kanehisa, *Nuc. Acids Res.*, 12:203-213 (1984), which is incorporated herein by reference. The length of homology comparison, as described, may be  
5 over longer stretches, and in certain embodiments will be over a stretch of at least about 17 nucleotides, generally at least about 20 nucleotides, ordinarily at least about 24 nucleotides, usually at least about 28 nucleotides, typically at least about 32 nucleotides,  
10 more typically at least about 40 nucleotides, preferably at least about 50 nucleotides, and more preferably at least about 75 to 100 or more nucleotides.

Amino acid sequence homology, or sequence  
15 identity, is determined by optimizing residue matches, if necessary, by introducing gaps as required. This changes when considering conservative substitutions as matches. Conservative substitutions typically include substitutions within the following groups: [glycine,  
20 alanine]; [valine, isoleucine, leucine]; [aspartic acid, glutamic acid]; [asparagine, glutamine]; [serine, threonine]; [lysine, arginine]; and [phenylalanine, tyrosine]. Homologous amino acid sequences are intended to include natural allelic and interspecies variations in  
25 each respective receptor sequence. Typical homologous proteins or peptides will have from 25-100% homology (if gaps can be introduced), to 50-100% homology (if conservative substitutions are included). Homology measures will be at least about 50%, generally at least  
30 56%, more generally at least 62%, often at least 67%, more often at least 72%, typically at least 77%, more typically at least 82%, usually at least 86%, more usually at least 90%, preferably at least 93%, and more

preferably at least 96%, and in particularly preferred embodiments, at least 98% or more.

In relation to proteins, the term "homology" in all its grammatical forms refers to the relationship between proteins that possess a "common evolutionary origin, " including proteins from superfamilies (e.g., the immunoglobulin superfamily) and homologous proteins from different species (e.g., myosin light chain, etc.) (Reeck et al., Cell, 50:667 (1987)). The present invention naturally contemplates homologues of the proteins disclosed herein, and polynucleotides encoding the same, as falling within the scope of the invention.

For sequence comparison, typically one sequence acts as a reference sequence, to which test sequences are compared. When using a sequence comparison algorithm, test and reference sequences are entered into a computer, subsequence coordinates are designated, if necessary, and sequence algorithm program parameters are designated. Default program parameters can be used, or alternative parameters can be designated. The sequence comparison algorithm then calculates the percent sequence identities for the test sequences relative to the reference sequence, based on the program parameters.

A "comparison window", as used herein, includes reference to a segment of any one of the number of contiguous positions selected from the group consisting of from 20 to 600, usually about 50 to about 200, more usually about 100 to about 150 in which a sequence may be compared to a reference sequence of the same number of contiguous positions after the two sequences are optimally aligned. Methods of alignment of sequences for

comparison are well-known in the art. Optimal alignment of sequences for comparison can be conducted, e.g., by the local homology algorithm of Smith & Waterman, *Adv. Appl. Math.* 2:482 (1981), by the homology alignment algorithm of Needleman & Wunsch, *J. Mol. Biol.* 48:443 (1970), by the search for similarity method of Pearson & Lipman, *Proc. Nat'l. Acad. Sci. USA* 85:2444 (1988), by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI), or by manual alignment and visual inspection (see, e.g., *Current Protocols in Molecular Biology* (Ausubel et al., eds. 1995 supplement)).

One example of a useful algorithm is PILEUP. PILEUP creates a multiple sequence alignment from a group of related sequences using progressive, pairwise alignments to show relationship and percent sequence identity. It also plots a tree or dendrogram showing the clustering relationships used to create the alignment. PILEUP uses a simplification of the progressive alignment method of Feng & Doolittle, *J. Mol. Evol.* 35:351-360 (1987). The method used is similar to the method described by Higgins & Sharp, *CABIOS* 5:151-153 (1989). The program can align up to 300 sequences, each of a maximum length of 5,000 nucleotides or amino acids. The multiple alignment procedure begins with the pairwise alignment of the two most similar sequences, producing a cluster of two aligned sequences. This cluster is then aligned to the next most related sequence or cluster of aligned sequences. Two clusters of sequences are aligned by a simple extension of the pairwise alignment of two individual sequences. The final alignment is achieved by a series of progressive, pairwise alignments. The

program is run by designating specific sequences and their amino acid or nucleotide coordinates for regions of sequence comparison and by designating the program parameters. Using PILEUP, a reference sequence is  
5 compared to other test sequences to determine the percent sequence identity relationship using the following parameters: default gap weight (3.00), default gap length weight (0.10), and weighted end gaps. PILEUP can be obtained from the GCG sequence analysis software package,  
10 e.g., version 7.0 (Devereaux et al., *Nuc. Acids Res.* 12:387-395 (1984).

Another example of algorithm that is suitable for determining percent sequence identity and sequence  
15 similarity are the BLAST and BLAST 2.0 algorithms, which are described in Altschul et al., *Nuc. Acids Res.* 25:3389-3402 (1977) and Altschul et al., *J. Mol. Biol.* 215:403-410 (1990), respectively. Software for performing BLAST analyses is publicly available through  
20 the National Center for Biotechnology Information (ncbi.nlm.nih.gov). This algorithm involves first identifying high scoring sequence pairs (HSPs) by identifying short words of length W in the query sequence, which either match or satisfy some positive-valued threshold score T when aligned with a word of the  
25 same length in a database sequence. T is referred to as the neighborhood word score threshold (Altschul et al., *supra*). These initial neighborhood word hits act as seeds for initiating searches to find longer HSPs  
30 containing them. The word hits are extended in both directions along each sequence for as far as the cumulative alignment score can be increased. Cumulative scores are calculated using, for nucleotide sequences, the parameters M (reward score for a pair of matching

residues; always  $> 0$ ) and N (penalty score for mismatching residues; always  $< 0$ ). For amino acid sequences, a scoring matrix is used to calculate the cumulative score. Extension of the word hits in each  
5 direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached.  
10 The BLAST algorithm parameters W, T, and X determine the sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength (W) of 11, an expectation (E) of 10, M=5, N=-4 and a comparison of both strands. For amino acid  
15 sequences, the BLASTP program uses as defaults a wordlength of 3, and expectation (E) of 10, and the BLOSUM62 scoring matrix (see Henikoff & Henikoff, *Proc. Natl. Acad. Sci. USA* 89:10915 (1989)) alignments (B) of 50, expectation (E) of 10, M=5, N=-4, and a comparison of  
20 both strands.

The BLAST algorithm also performs a statistical analysis of the similarity between two sequences (see, e.g., Karlin & Altschul, *Proc. Nat'l. Acad. Sci. USA*  
25 90:5873-5787 (1993)). One measure of similarity provided by the BLAST algorithm is the smallest sum probability (P(N)), which provides an indication of the probability by which a match between two nucleotide or amino acid sequences would occur by chance. For example, a nucleic  
30 acid is considered similar to a reference sequence if the smallest sum probability in a comparison of the test nucleic acid to the reference nucleic acid is less than about 0.2, more preferably less than about 0.01, and most preferably less than about 0.001.



An indication that two nucleic acid sequences or polypeptides are substantially identical is that the polypeptide encoded by the first nucleic acid is  
5 immunologically cross reactive with the antibodies raised against the polypeptide encoded by the second nucleic acid, as described below. Thus, a polypeptide is typically substantially identical to a second polypeptide, for example, where the two peptides differ  
10 only by conservative substitutions. Another indication that two nucleic acid sequences are substantially identical is that the two molecules or their complements hybridize to each other under stringent conditions, as described below. Yet another indication that two nucleic  
15 acid sequences are substantially identical is that the same primers can be used to amplify the sequence.

It should be understood that each method of the invention described herein encompasses: a) all compounds  
20 having about 80%, 85%, 90%, 92%, 95%, 96%, 97%, 98% or 99% identity with SEQ ID NOS:1 to 12 and RNA correlates of SEQ ID NOS:1 to 6; b) all compounds with 50, 60, 70, 80, 90 or more amino acids and having about 92%, 95%, 96%, 97%, 98% or 99% identity with SEQ ID NOS:7 to 12; c)  
25 all compounds having about 80%, 85%, 90%, 92%, 95%, 96%, 97%, 98% or 99% identity with all domains of SEQ ID NOS:6 to 10; and c) all compounds with 100, 150, 200, 250, 300 or more nucleotides and having about 92%, 95%, 96%, 97%, 98% or 99% identity with SEQ ID NOS:1 to 6 or their RNA  
30 correlates. The invention also encompasses all encoding DNA of SEQ ID NOS:7 to 12, as well as RNA correlates of such DNA.

The tem "moderately stringent conditions," as used herein, means hybridization conditions that permit a nucleic acid molecule to bind to a second nucleic acid molecule that has substantial identity to the sequence of the first. Moderately stringent conditions are those equivalent to hybridization of filer-bound nucleic acid in 50% formamide, 5 X Denhart's solution, 5 X SSPE, 0.2% SDS at 42°C, followed by washing in 0.2 X SSPE, 0.2% SDS at 50°C. "Highly stringent conditions" are those equivalent to hybridization of filer-bound nucleic acid in 50% formamide, 5 X Denhart's solution, 5 X SSPE, 0.2% SDS at 42°C, followed by washing in 0.2 X SSPE, 0.2% SDS at 65°C. Other suitable moderately stringent and highly stringent conditions are known in the art and described, for example, in Sambrook et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, New York (1992), and Ansubel et al., Current Protocols in Molecular Biology, John Wiley and Sons, Baltimore MD (1998).

20

In general, a nucleic acid molecule that hybridizes to a second one under moderately stringent conditions will have greater than about 60% identity, preferably greater than about 70% identity and, more preferably, greater than about 80% identity over the length of the two sequences being compared. A nucleic acid molecule that hybridizes to a second one under highly stringent conditions will have greater than about 90% identity, preferably greater than about 92% identity and, more preferably, greater than about 95%, 96%, 97%, 98% or 99% identity over the length of the two sequences being compared.

30

As used herein, the term "isolated" when used in conjunction with a nucleic acid or protein, denotes that the nucleic acid or protein has been isolated with respect to the many other cellular components with which it is normally associated in the natural state. For example, an "isolated" gene of interest may be one that has been separated from open reading frames which flank the gene and encode a gene product other than that of the specific gene of interest. Such genes may be obtained by a number of methods including, for example, laboratory synthesis, restriction enzyme digestion or PCR. Likewise, an "isolated" protein may be substantially purified from a natural source or may be synthesized in the laboratory. A "substantially purified" nucleic acid or protein gives rise to essentially one band in an electrophoretic gel, and is at least 85% pure, more preferably at least 95% pure, and most preferably at least 99% pure.

The term "intrabody" refers to a class of neutralizing molecules with applications in gene therapy (vonMehren M, Weiner L M. (1996) *Current Opinion in Oncology*. 8:493-498, Marasco Wash. (1997) *Gene Therapy*. 4:11-15, Rondon I J, Marasco Wash. (1997) *Annual Review of Microbiology*. 51:257-283). Intrabodies are engineered antibodies that can be expressed within a cell and target an intracellular molecule or molecular domain. Using this technique, intracellular signals and enzyme activities can be inhibited, or their transport to cellular compartments prevented. Marasco, W. A., et al., *Proc. Natl. Acad. Sci. USA* 90:7889-7893 (1993). Thus, intrabodies provide yet another approach to down regulating protein expression and activity.

The intrabody method is analogous to the inactivation of proteins by deletion or mutation, but is directed at the level of gene product rather than at the gene itself. Using the intrabody strategy even molecules  
5 involved in essential cellular pathways can be targeted, modified or blocked. Antibody genes for intracellular expression can be derived either from murine or human monoclonal antibodies or from phage display libraries. For intracellular expression small recombinant antibody  
10 fragments, containing the antigen recognizing and binding regions, can be used. Intrabodies can be directed to different intracellular compartments by targeting sequences attached to the antibody fragments. The construction and use of intrabodies is discussed, for  
15 example, in U.S. Pat. No. 6,004,940.

As used herein, the term "expression vector" includes a recombinant expression cassette that has a nucleotide sequence that can be transcribed into RNA in a  
20 cell. The cell can further translate transcribed mRNA into protein. An expression vector can be a plasmid, virus, or nucleic acid fragment. Typically, the recombinant expression cassette portion of an expression vector includes the encoding nucleotide sequence to be  
25 transcribed (e.g. a ribozyme or siRNA), operably linked to a promoter, or other regulatory sequence by a functional linkage in cis. In accordance with the present invention, an expression vector comprising a nucleotide sequence encoding ribozymes of the invention  
30 can be used to transduce cells suitable as hosts for the vector. Both procaryotic cells including bacterial cells such as *E. coli* and eukaryotic cells including mammalian cells may be used for this purpose.

As used herein, the term "promoter" includes nucleic acid sequences near the start site of transcription (such as a polymerase binding site) and, optionally, distal enhancer or repressor elements (which  
5 may be located several thousand base pairs from the start site of transcription) that direct transcription of the nucleotide sequence in a cell. The term includes both a "constitutive" promoter such as a pol III promoter, which is active under most environmental conditions and stages  
10 of development or cell differentiation, and an "inducible" promoter, which initiates transcription in response to an extracellular stimulus, such as a particular temperature shift or exposure to a specific chemical. Promoters and other regulatory elements (e.g.,  
15 an origin of replication), and/or chromosome integration elements such as retroviral long terminal repeats ("LTRs"), or adeno associated viral (AAV) inverted terminal repeats ("ITRs"), may be incorporated into an expression vector encoding ribozymes of the present  
20 invention as described in WO 00/05415 to Barber *et al.*

As used herein, the term "expresses" denotes that a given nucleic acid comprising an open reading frame is transcribed to produce an RNA molecule. It also  
25 denotes that a given nucleic acid is transcribed and translated to produce a polypeptide. Although the term may be used to refer to the transcription of a ribozyme, a ribozyme typically is not translated into a protein since it functions as an active (catalytic) nucleic acid.

30

As used herein, the term "gene product" refers either to the RNA produced by transcription of a given nucleic acid or to the polypeptide produced by translation of a given nucleic acid.

As used herein, the term "transduce" denotes the introduction of an exogenous nucleic acid molecule (e.g., by means of an expression vector) inside the membrane of a cell. Exogenous DNA may or may not be integrated (covalently linked) into chromosomal DNA making up the genome of the cell. The exogenous DNA may be maintained on an episomal element, such as a plasmid. In eukaryotic cells, a stably transduced cell is generally one in which the exogenous DNA has become integrated into the chromosome so that it is inherited by daughter cells through chromosome replication, or one which includes stably maintained extrachromosomal plasmids. This stability is demonstrated by the ability of the eukaryotic cell to establish cell lines or clones comprised of a population of daughter cells containing the exogenous DNA.

The term "transfection," as used herein, means the genetic modification of a cell by uptake of an exogenous nucleic acid molecule (e.g., by means of an expression vector).

As used herein, the term "ribozyme gene vector library" denotes a collection of ribozyme-encoding genes, typically within expression cassettes, in a collection of viral or other vectors. The vectors may be naked or contained within a capsid. Propagation of the ribozyme gene vector library can be performed as described in WO 00/05415 to Barber et al. The ribozyme-encoding genes of a ribozyme gene vector library, after transduction and transcription in appropriate cells, produce a collection of ribozymes.

As used herein, "small interfering RNAs" (siRNA) are short double-stranded RNA fragments that elicit a process known as "RNA interference" (RNAi), a form of

5 sequence-specific gene silencing. Zamore, Phillip et al., Cell 101:25-33(2000); Elbashir, Sayda M., et al., Nature 411:494-497 (2001). siRNAs are assembled into a multi-component complex known as the RNA-induced silencing complex (RISC). The siRNAs guide RISC to

10 homologous mRNAs, thus targeting them for destruction. Hammond et al., Nature Genetics Reviews 2:110-119(2000). RNAi has been observed in a variety of organisms including plants, insects and mammals, and cultured cells derived from these organisms.

15

An "siRNA" is a double-stranded RNA that is preferably between 16 and 25, more preferably 17 and 23 and most preferably between 18 and 21 base pairs long, each strand of which has a 3' overhang of 2 or more

20 nucleotides. Functionally, the characteristic distinguishing an siRNA over other forms of dsRNA is that the siRNA comprises a sequence capable of specifically inhibiting genetic expression of a gene or closely related family of genes by a process termed RNA

25 interference.

siRNAs for use in the present invention can be produced from a gene of the invention (SEQ ID NOS:1 to 6). For example, short complementary DNA strands are

30 first prepared that represent portions of both the "sense" and "antisense" strands of the NR4A1 coding region. This is typically accomplished using solid phase nucleic acid synthesis techniques, as known in the art. The short duplex DNA thus formed is ligated into a

suitable vector that is then used to transfect a suitable cell line. Other methods for producing siRNA molecules are known in the art. (See, e.g., Elbashir, S.M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K., and Tuschl, T. (2001). For a review of RNAi and siRNA expression, see Hammond, Scott M. et al., *Nature Genetics Reviews*, 2:110-119; Fire, Andrew (1999) *TIG*, 15(9):358-363; Bass, Brenda L. (2000) *Cell*, 101:235-238. An siRNA of the invention can be constructed using, for example, a Lentiviral vector for stable expression.

siRNA molecules can be transfected into a cell line, for example HeLa, by using an agent such as Oligofectamine<sup>TM</sup>, as described in Example 4. See also Invitrogen Corp., Transfecting siRNA into HeLa Cells Using Oligofectamine<sup>TM</sup>, Doc. Rev. 102902 (Carlsbad, CA); Elbashir, et al., *Nature*, 411:494-498 (2001); and Harborth et al., *Science*, 114:4557-4565 (2001).

The targeting of antisense oligonucleotides to mRNA is another mechanism of decreasing protein synthesis, and, consequently, represents a powerful and targeted approach to diminishing expression of the proteins of the invention. For example, the synthesis of polygalacturonase and the muscarine type 2 acetylcholine receptor are inhibited by antisense oligonucleotides directed to their respective mRNA sequences (U.S. Pat. Nos. 5,739,119 and 5,759,829, each specifically incorporated herein by reference in its entirety). Further, examples of antisense inhibition have been demonstrated with the nuclear protein cyclin, the multiple drug resistance gene (MDG1), ICAM-1, E-selectin, STK-1, striatal GABA.sub.A receptor and human EGF (Jaskulski et al., 1988; Vasanthakumar and Ahmed, 1989;



Peris et al., 1998; U.S. Pat. Nos. 5,801,154; 5,789,573; 5,718,709 and 5,610,288, each specifically incorporated herein by reference in its entirety). Antisense constructs have also been described that inhibit and can  
5 be used to treat a variety of abnormal cellular proliferations, e.g. cancer (U.S. Pat. Nos. 5,747,470; 5,591,317 and 5,783,683, each specifically incorporated herein by reference in its entirety).

10           The invention provides therefore oligonucleotide sequences that comprise all, or a portion of, any sequence that is capable of specifically binding to a polynucleotide sequence described herein, or a complement thereof. In one embodiment, the antisense  
15 oligonucleotides comprise DNA or derivatives thereof. In another embodiment, the oligonucleotides comprise RNA or derivatives thereof. In a third embodiment, the oligonucleotides are modified DNAs comprising a phosphorothioated modified backbone. In a fourth  
20 embodiment, the oligonucleotide sequences comprise peptide nucleic acids or derivatives thereof. In each case, preferred compositions comprise a sequence region that is complementary, and more preferably substantially-complementary, and even more preferably, completely  
25 complementary to one or more portions of SEQ ID NOS:1, 2, 3, 4 or 5.

          Selection of antisense compositions specific for a given gene sequence is based upon analysis of the  
30 chosen target sequence and determination of secondary structure, binding energy, and relative stability. Antisense compositions are selected based upon their relative inability to form dimers, hairpins, or other

secondary structures that would reduce or prohibit specific binding to the target mRNA in a host cell.

"Non-native promoter" refers to any promoter element operably linked to a coding sequence by recombinant methods. Non-native promoters include mutagenized native reporters, when mutagenesis alters the rate or control of transcriptional events.

"Operably linked" refers to a linkage of polynucleotide elements in a functional relationship. With regard to the present invention, the term "operably linked" refers to a functional linkage between a nucleic acid expression control sequence (such as a promoter, or an array of transcription factor binding sites) and a second nucleic acid sequence, wherein the expression control sequence directs transcription of the nucleic acid corresponding to the second sequence. Thus, a nucleic acid is "operably linked" when it is placed into a functional relationship with another nucleic acid sequence.

The invention also encompasses vectors in which a nucleic acid is cloned into a vector in reverse orientation, but operably linked to a regulatory sequence that permits transcription of antisense RNA. Thus, an antisense transcript can be produced to all, or to a portion, of the nucleic acid sequences described herein, including both coding and non-coding regions. Expression of this antisense RNA is subject to each of the parameters described above in relation to expression of the sense RNA (regulatory sequences, constitutive or inducible expression, tissue-specific expression).

Antisense nucleic acids may be obtained from libraries encoding proteins of the invention (SEQ ID NOS:7 to 12) or synthesized synthetically. Transfection of suitable host cells with such a protein is performed  
5 in a manner analogous to that described for siRNAs above.

"Recombinant expression cassette" refers to a DNA sequence capable of directing expression of a nucleic acid in cells. A "DNA expression cassette" comprises a  
10 promoter, operably linked to a nucleic acid of interest, which is further operably linked to a termination region.

Where a range of values is provided, it is understood that each intervening value, to the tenth of  
15 the unit of the lower limit unless the context clearly dictates otherwise, between the upper and lower limits of that range is also specifically disclosed. Each smaller range between any stated value or intervening value in a stated range and any other stated or intervening value in  
20 that stated range is encompassed within the invention. The upper and lower limits of these smaller ranges may independently be included or excluded in the range, and each range where either, neither or both limits are included in the smaller ranges is also encompassed within  
25 the invention, subject to any specifically excluded limit in the stated range. Where the stated range includes one or both of the limits, ranges excluding either or both of those included limits are also included in the invention.

30 Polypeptides often contain amino acids other than the 20 amino acids commonly referred to as the 20 naturally-occurring amino acids. Further, many amino acids, including the terminal amino acids, may be modified by natural processes, such as processing and

other post-translational modifications, or by chemical modification techniques well known in the art. Common modifications that occur naturally in polypeptides are described in basic texts, detailed monographs, and the  
5 research literature, and they are well known to those of skill in the art.

Accordingly, the polypeptides also encompass derivatives or analogs in which a substituted amino acid  
10 residue is not one encoded by the genetic code, in which a substituent group is included, in which the mature polypeptide is fused with another compound, such as a compound to increase the half-life of the polypeptide (for example, polyethylene glycol), or in which the  
15 additional amino acids are fused to the mature polypeptide, such as a leader or secretory sequence or a sequence for purification of the mature polypeptide or a pro-protein sequence.

20 Known modifications include, but are not limited to, acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment  
25 of a lipid or lipid derivative, covalent attachment of phosphatidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent crosslinks, formation of cystine, formation of pyroglutamate, formylation, gamma carboxylation,  
30 glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, transfer-RNA

mediated addition of amino acids to proteins such as arginylation, and ubiquitination.

Such modifications are well-known to those of skill in the art and have been described in great detail in the scientific literature. Several particularly common modifications, glycosylation, lipid attachment, sulfation, gamma-carboxylation of glutamic acid residues, hydroxylation and ADP-ribosylation, for instance, are described in most basic texts, such as *Proteins--Structure and Molecular Properties*, 2nd Ed., T. E. Creighton, W. H. Freeman and Company, New York (1993). Many detailed reviews are available on this subject, such as by Wold, F., *Posttranslational Covalent Modification of Proteins*, B. C. Johnson, Ed., Academic Press, New York 1-12 (1983); Seifter et al. (*Meth. Enzymol.* **182**: 626-646 (1990)) and Rattan et al. (*Ann. N.Y. Acad. Sci.* **663**:48-62 (1992)).

As is also well known, polypeptides are not always entirely linear. For instance, polypeptides may be branched as a result of ubiquitination, and they may be circular, with or without branching, generally as a result of post-translation events, including natural processing event and events brought about by human manipulation which do not occur naturally. Circular, branched and branched circular polypeptides may be synthesized by non-translational natural processes and by synthetic methods.

30

Modifications can occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. Blockage of the amino or carboxyl group in a polypeptide,

or both, by a covalent modification, is common in naturally-occurring and synthetic polypeptides. For instance, the amino terminal residue of polypeptides made in *E. coli*, prior to proteolytic processing, almost  
5 invariably will be N-formylmethionine.

The modifications can be a function of how the protein is made. For recombinant polypeptides, for example, the modifications will be determined by the host  
10 cell posttranslational modification capacity and the modification signals in the polypeptide amino acid sequence. Accordingly, when glycosylation is desired, a polypeptide should be expressed in a glycosylating host, generally a eukaryotic cell. Insect cells often carry  
15 out the same posttranslational glycosylations as mammalian cells and, for this reason, insect cell expression systems have been developed to efficiently express mammalian proteins having native patterns of glycosylation. Similar considerations apply to other  
20 modifications.

The same type of modification may be present in the same or varying degree at several sites in a given polypeptide. Also, a given polypeptide may contain more  
25 than one type of modification.

Candidate protein-based compounds for binding or down-modulating a protein of the invention or one of its ZF or HOLI domains include, for example, 1) peptides  
30 such as soluble peptides, including fusion peptides and members of random peptide libraries (see, e.g., Lam et al., *Nature* 354:82-84 (1991); Houghten et al., *Nature* 354:84-86 (1991)) and combinatorial chemistry-derived molecular libraries made of L- and/or D-configuration

amino acids; 2) phosphopeptides (e.g., members of random and partially degenerate, directed phosphopeptide libraries, see, e.g., Songyang et al., *Cell* 72:767-778 (1993)); 3) antibodies (e.g., polyclonal, monoclonal, 5 humanized, anti-idiotypic, chimeric, and single chain antibodies, including intrabodies, as well as Fab, F(ab')<sub>2</sub>, Fab expression library fragments, and epitope-binding fragments of antibodies); and 4) small organic and inorganic molecules (e.g., molecules obtained 10 from combinatorial and natural product libraries).

Soluble full-length receptors, or fragments of the same, that compete for ligand binding are also considered candidate reagents. Other candidate compounds 15 include mutant receptors or appropriate fragments containing mutations that affect receptor function and thus compete for ligand. Accordingly, a fragment that competes for ligand, for example with a higher affinity, or a fragment that binds ligand but does not allow 20 release, is encompassed by the invention. The receptor polynucleotides are also useful for constructing host cells expressing a part, or all, of the receptor polynucleotides and polypeptides.

25 The receptor polynucleotides are also useful for constructing transgenic animals expressing all, or a part, of the receptor polynucleotides and polypeptides. These animals are useful as model systems for the treatment of cancer and can be used to test compounds for 30 their effect, through the receptor gene or gene product, on the development or progression of the disease.

As used herein, the term "knockdown" or "down-modulation" means a decrease in the rate or level of mRNA

production and/or protein production. For example, siRNA constructs of the present invention were shown to down-modulate the mRNA of the target genes of the invention (SEQ ID NOS:1 to 6), as discussed above and in the  
5 examples below.

The present invention also provides a method of inhibiting cancer cells or their growth. This method includes introducing an agent of the invention into  
10 cancer cells. Such an agent can be, for example, a ribozyme of the present invention, an antisense molecule, an antagonizing antibody or an siRNA. This method can comprise transducing the infected cell with an expression vector encoding the agent. Alternatively, the agent can  
15 be introduced into a cell directly, i.e., without using a vector.

The method of the invention can be accomplished by the agent binding to a target, for instance SEQ ID  
20 NOS: 1 to 10, or portions or domains thereof, as discussed above. The method of the invention can also be accomplished by the agent down-modulating SEQ ID NOS:1 to 10, or portions or domains thereof, as discussed above.

25 In accordance with another embodiment of the present invention, the step of down-modulating the level of the target protein in the cell can be accomplished by introducing into the cell system an antisense compound or molecule.

30 Combinatorial peptide libraries can be screened to identify antagonists of a protein, ZF domain or HOLI domain of the invention, which can increase the inhibition of cancer cells. Combinatorial peptide



libraries can be constructed from genomic or cDNA libraries, or by using non-cellular synthetic methods. Techniques for solid phase synthesis are described by Barany and Merrifield, *Solid-Phase Peptide Synthesis*, pp. 3-284 in *The Peptides: Analysis, Synthesis, Biology*. Vol. 2: *Special Methods in Peptide Synthesis, Part A.*; Merrifield, et al., *J. Am. Chem. Soc.* **85**: 2149-2156 (1963), and Stewart et al., *Solid Phase Peptide Synthesis*, 2nd ed., Pierce Chem. Co., Rockford, III. (1984). Proteins may be synthesized by condensation of the amino and carboxy termini of shorter fragments. Methods of forming peptide bonds by activation of a carboxy terminal end (e.g., by the use of the coupling reagent N,N'-dicyclohexylcarbodiimide) are known to those of skill.

The proteins useful in this invention may be purified to substantial purity by standard techniques well known in the art, including detergent solubilization, selective precipitation with such substances as ammonium sulfate, column chromatography, immunopurification methods, and others. See, for instance, R. Scopes, *Protein Purification: Principles and Practice*, Springer-Verlag: New York (1982); Deutscher, *Guide to Protein Purification*, Academic Press (1990). For example, antibodies may be raised to the proteins as described herein. Purification from *E. coli* can be achieved following procedures described in U.S. Pat. No. 4,511,503.

Peptide and protein reagents can optionally be labeled, as described below, or may be used in the screening assays of the present invention to ascertain their ability to modulate protein expression or activity.

Portions of SEQ ID NOS:7 to 12 can be useful in competition binding assays in methods designed to discover compounds that interact with the receptor.

5 Thus, a compound is exposed to a receptor polypeptide under conditions that allow the compound to bind or to otherwise interact with the polypeptide. Soluble receptor polypeptide is also added to the mixture. If the test compound interacts with the soluble receptor  
10 polypeptide, it decreases the amount of complex formed or activity from the receptor target. This type of assay is particularly useful in cases in which compounds are sought that interact with specific regions of the receptor. Thus, the soluble polypeptide that competes  
15 with the target receptor region is designed to contain peptide sequences corresponding to the region of interest.

The compounds tested as modulators of proteins  
20 or protein domains of the invention or RNA correlates of the invention can be any small chemical compound, or a biological entity, such as a protein, sugar, nucleic acid or lipid. Screening combinatorial libraries of small organic molecules offers an approach to identifying  
25 useful therapeutic compounds or precursors targeted to proteins or protein domains of the invention or RNA correlates of the invention. Typically, test compounds will be small chemical molecules and peptides. Essentially any chemical compound can be used as a  
30 potential modulator or ligand in the assays of the invention, although most often compounds can be dissolved in aqueous or organic (especially DMSO-based) solutions are used. The assays are designed to screen large chemical libraries by automating the assay steps and

providing compounds from any convenient source to assays, which are typically run in parallel (e.g., in microtiter formats on microtiter plates in robotic assays). It will be appreciated that there are many suppliers of chemical  
5 compounds, including Sigma (St. Louis, MO), Aldrich (St. Louis, MO), Sigma-Aldrich (St. Louis, MO), Fluka Chemika-Biochemica Analytika (Buchs Switzerland) and the like.

In one embodiment, high throughput screening  
10 methods are utilized involving a combinatorial chemical or peptide library containing a large number of potential therapeutic compounds (potential modulator or ligand compounds). Such "combinatorial chemical libraries" or "ligand libraries" are screened in one or more assays, as  
15 described herein, to identify those library members (particular chemical species or subclasses) that display a desired characteristic activity. The compounds thus identified can serve as conventional "lead compounds" or can themselves be used as potential or actual  
20 therapeutics.

A combinatorial chemical library is a collection of diverse chemical compounds generated by either chemical synthesis or biological synthesis, by  
25 combining a number of chemical "building blocks" such as reagents. For example, a linear combinatorial chemical library such as a polypeptide library is formed by combining a set of chemical building blocks (amino acids) in every possible way for a given compound length (i.e.,  
30 the number of amino acids in a polypeptide compound). Millions of chemical compounds can be synthesized through such combinatorial mixing of chemical building blocks. Preparation and screening of combinatorial chemical libraries is well known to those of skill in the art.

To perform cell-free drug screening assays, it is desirable to immobilize either the receptor protein, or fragment, or its target molecule to facilitate  
5 separation of complexes from uncomplexed forms of one or both of the proteins, as well as to accommodate automation of the assay.

In one embodiment, the invention provides  
10 soluble assays using molecules such as a ligand binding domain, an extracellular domain, a transmembrane domain (e.g., one comprising seven transmembrane regions and cytosolic loops), the transmembrane domain and a cytoplasmic domain, an active site, a subunit association  
15 region, etc.; a domain that is covalently linked to a heterologous protein to create a chimeric molecule; a protein of the invention; or a cell or tissue expressing such protein, either naturally occurring or recombinant. In another embodiment, the invention provides solid phase  
20 based *in vitro* assays in a high throughput format, where the domain, chimeric molecule, protein of the invention, or cell or tissue expressing such protein is attached to a solid phase substrate.

25 In the high throughput assays of the invention, it is possible to screen up to several thousand different modulators or ligands in a single day. In particular, each well of a microtiter plate can be used to run a separate assay against a selected potential modulator,  
30 or, if concentration or incubation time effects are to be observed, every 5-10 wells can test a single modulator. Thus, a single standard microtiter plate can assay about 100 (e.g., 96) modulators. If 1536 well plates are used, then a single plate can easily assay from about 100-

[????] about 1500 different compounds. It is possible to assay several different plates per day; assay screens for up to about 6,000-20,000 different compounds is possible using the integrated systems of the invention. More recently, microfluidic approaches to reagent manipulation have been developed, e.g., by Caliper Technologies (Palo Alto, CA).

The molecule of interest can be bound to the solid state component, directly or indirectly, via covalent or non covalent linkage e.g., via a tag. The tag can be any of a variety of components. In general, a molecule which binds the tag (a tag binder) is fixed to a solid support, and the tagged molecule of interest (e.g., the taste transduction molecule of interest)[????] is attached to the solid support by interaction of the tag and the tag binder.

A number of tags and tag binders can be used, based upon known molecular interactions well described in the literature. For example, where a tag has a natural binder, for example, biotin, protein A, or protein G, it can be used in conjunction with appropriate tag binders (avidin, streptavidin, neutravidin, the Fc region of an immunoglobulin, etc.) Antibodies to molecules with natural binders such as biotin are also widely available as are appropriate tag binders; see, SIGMA Immunochemicals 1998 catalogue SIGMA, St. Louis MO).

Similarly, any haptenic or antigenic compound can be used in combination with an appropriate antibody to form a tag/tag binder pair. Thousands of specific antibodies are commercially available and many additional antibodies are described in the literature. For example,

in one common configuration, the tag is a first antibody and the tag binder is a second antibody that recognizes the first antibody. In addition to antibody-antigen interactions, receptor-ligand interactions are also  
5 appropriate as tag and tag-binder pairs. For example, agonists and antagonists of cell membrane receptors (e.g., cell receptor-ligand interactions such as transferrin, c-kit, viral receptor ligands, cytokine receptors, chemokine receptors, interleukin receptors,  
10 immunoglobulin receptors and antibodies, the cadherein family, the integrin family, the selectin family, and the like; see, e.g., Pigott & Power, *The Adhesion Molecule Facts Book I* (1993). Similarly, toxins and venoms, viral epitopes, hormones (e.g., opiates, steroids, etc.),  
15 intracellular receptors (e.g. which mediate the effects of various small ligands, including steroids, thyroid hormone, retinoids and vitamin D; peptides), drugs, lectins, sugars, nucleic acids (both linear and cyclic polymer configurations), oligosaccharides, proteins,  
20 phospholipids and antibodies can all interact with various cell receptors.

Synthetic polymers, such as polyurethanes, polyesters, polycarbonates, polyureas, polyamides,  
25 polyethyleneimines, polyarylene sulfides, polysiloxanes, polyimides, and polyacetates can also form an appropriate tag or tag binder. Many other tag/tag binder pairs are also useful in assay systems described herein, as would be apparent to one of skill upon review of this  
30 disclosure.

Common linkers such as peptides, polyethers, and the like can also serve as tags, and include polypeptide sequences, such as poly gly sequences of

between about 5 and 200 amino acids. Such flexible linkers are known to persons of skill in the art. For example, poly(ethelyne glycol) linkers are available from Shearwater Polymers, Inc. Huntsville, Alabama. These  
5 linkers optionally have amide linkages, sulfhydryl linkages, or heterofunctional linkages.

Tag binders are fixed to solid substrates using any of a variety of methods currently available. Solid  
10 substrates are commonly derivatized or functionalized by exposing all or a portion of the substrate to a chemical reagent which fixes a chemical group to the surface which is reactive with a portion of the tag binder. For example, groups which are suitable for attachment to a  
15 longer chain portion would include amines, hydroxyl, thiol, and carboxyl groups. Aminoalkylsilanes and hydroxyalkylsilanes can be used to functionalize a variety of surfaces, such as glass surfaces. The construction of such solid phase biopolymer arrays is  
20 well described in the literature. See, e.g., Merrifield, *J. Am. Chem. Soc.* 85:2149-2154 (1963) (describing solid phase synthesis of, e.g., peptides); Geysen et al., *J. Immun. Meth.* 102:259-274 (1987) (describing synthesis of solid phase components on pins); Frank & Doring,  
25 *Tetrahedron* 44:60316040 (1988) (describing synthesis of various peptide sequences on cellulose disks); Fodor et al., *Science*, 251:767-777 (1991); Sheldon et al., *Clinical Chemistry* 39(4):718-719 (1993); and Kozal et al., *Nature Medicine* 2(7):753759 (1996) (all describing  
30 arrays of biopolymers fixed to solid substrates). Non-chemical approaches for fixing tag binders to substrates include other common methods, such as heat, cross-linking by UV radiation, and the like.

Another approach uses recombinant bacteriophage to produce large libraries. Using the "phage method" (Scott and Smith, *Science* 249:386-390, 1990; Cwirla, et al, *Proc. Natl. Acad. Sci.*, 87:6378-6382, 1990; Devlin et al., *Science*, 49:404-406, 1990), very large libraries can be constructed ( $10^6$  - $10^8$  chemical entities). A second approach uses primarily chemical methods, of which the Geysen method (Geysen et al., *Molecular Immunology* 23:709-715, 1986; Geysen et al. *J. Immunologic Method* 102:259-274, 1987; and the method of Fodor et al. (*Science* 251:767-773, 1991) are examples. Furka et al. (14th International Congress of Biochemistry, Volume #5, Abstract FR:013, 1988; Furka, *Int. J. Peptide Protein Res.* 37:487-493, 1991), Houghton (U.S. Pat. No. 4,631,211, issued December 1986) and Rutter et al. (U.S. Pat. No. 5,010,175, issued Apr. 23, 1991) describe methods to produce a mixture of peptides that can be tested as agonists or antagonists.

In another aspect, synthetic libraries (Needels et al, *Proc. Natl. Acad. Sci. USA* 90:10700-4, 1993; Ohlmeyer et al., *Proc. Natl. Acad. Sci. USA* 90:10922-10926, 1993; Lam et al., International Patent Publication No. WO 92/00252; Kocis et al., International Patent Publication No. WO 9428028) and the like can be used to screen for a protein or domain according to the present invention.

The screening can be performed with recombinant cells that express the protein or domain of the invention, or alternatively, using purified protein, e.g., produced recombinantly, as described above. For example, the ability of labeled, soluble or solubilized



NR4A1 that includes the ligand-binding portion of the molecule, to bind ligand can be used.

Radioligand binding assays allow further  
5 characterization of hits from high throughput screens as well as analogs of neurotensin agonists and antagonists. Using membranes from cells stably expressing each neurotensin receptor subtype, one point binding assays are first performed to determine how well a particular  
10 concentration of each hit or analog displaces specific [<sup>3</sup>H] NT binding from the receptor. If the hit or analog displaces 50% of the [<sup>3</sup>H] NT bound, a competition binding assay is performed. Competition binding assays can evaluate the ability of increasing concentrations of  
15 competitor (the hit or any test compound analog) to displace [<sup>3</sup>H] NT binding at each neurotensin receptor subtype. The resulting K<sub>i</sub> value indicates the relative potency of each hit or test compound for a particular receptor subtype. These competition binding assays allow  
20 the determination of the relative potencies of each hit or test compound at a particular receptor subtype, as well as to determine the receptor subtype selectivity of each hit or test compound.

25 Yet another assay for compounds that modulate protein activity involves computer assisted drug design, in which a computer system is used to generate a three-dimensional structure of a protein of the invention based on the structural information encoded by the amino acid  
30 sequence. The input amino acid sequence interacts directly and actively with a pre-established algorithm in a computer program to yield secondary, tertiary, and quaternary structural models of the protein. The models of the protein structure are then examined to identify

regions of the structure that have the ability to bind, e.g., ligands. These regions are then used to identify ligands that bind to the protein.

5           The three-dimensional structural model of the protein is generated by entering protein amino acid sequences of at least 10 amino acid residues or corresponding nucleic acid sequences encoding a polypeptide into the computer system. Contiguous  
10 portions of SEQ ID NOS:1 to 6, and conservatively modified versions thereof, can be used for this purpose. The amino acid sequence represents the primary sequence or subsequence of the protein, which encodes the structural information of the protein. At least 10  
15 residues of the amino acid sequence (or a nucleotide sequence encoding 10 amino acids) are entered into the computer system from computer keyboards, computer readable substrates that include, but are not limited to, electronic storage media (e.g., magnetic diskettes,  
20 tapes, cartridges, and chips), optical media (e.g., CD ROM), information distributed by internet sites, and by RAM. The three-dimensional structural model of the protein is then generated by the interaction of the amino acid sequence and the computer system, using software  
25 known to those of skill in the art.

          The amino acid sequence represents a primary structure that encodes the information necessary to form the secondary, tertiary and quaternary structure of the  
30 protein of interest. The software looks at certain parameters encoded by the primary sequence to generate the structural model. These parameters are referred to as "energy terms," and primarily include electrostatic potentials, hydrophobic potentials, solvent accessible

surfaces, and hydrogen bonding. Secondary energy terms include van der Waals potentials. Biological molecules form the structures that minimize the energy terms in a cumulative fashion. The computer program therefore uses  
5 these terms encoded by the primary structure or amino acid sequence to create the secondary structural model.

The tertiary structure of the protein encoded by the secondary structure is then formed on the basis of  
10 the energy terms of the secondary structure. The user at this point can enter additional variables such as whether the protein is membrane bound or soluble, its location in the body, and its cellular location, e.g., cytoplasmic, surface, or nuclear. These variables along with the  
15 energy terms of the secondary structure are used to form the model of the tertiary structure. In modeling the tertiary structure, the computer program matches hydrophobic faces of secondary structure with like, and hydrophilic faces of secondary structure with like.

20

Once the structure has been generated, potential ligand binding regions are identified by the computer system. Three-dimensional structures for potential ligands are generated by entering amino acid or  
25 nucleotide sequences or chemical formulas of compounds, as described above. The three-dimensional structure of the potential ligand is then compared to that of the protein of the invention to identify ligands that bind to such protein. Binding affinity between the protein and  
30 ligands is determined using energy terms to determine which ligands have an enhanced probability of binding to the protein.

The activity of polypeptides can be assessed using a variety of *in vitro* and *in vivo* assays that determine functional, physical and chemical effects, e.g., measuring ligand binding (e.g., by radioactive  
5 ligand binding), second messengers (e.g., cAMP, cGMP, IP<sub>3</sub>, DAG, or Ca<sup>2+</sup>), ion flux, phosphorylation levels, transcription levels, neurotransmitter levels, and the like. Furthermore, such assays can be used to test for inhibitors and activators of proteins or domains of the  
10 invention. Modulators can also be genetically altered versions of proteins of the invention. Such modulators can be useful in the treatment and diagnosis of cancer.

The polypeptide of the assay will be selected  
15 from SEQ ID NOS:7 to 12, a portion of 10, 20, 30 , 40, 50 or more contiguous amino acids thereof, or conservatively modified variants thereof. Alternatively, the protein of the assay will be derived from a eukaryote and include an amino acid subsequence where the homology will be at  
20 least 60%, preferably at least 75%, more preferably at least 90% and most preferably between 95% and 100% that of SEQ ID NOS:7 to 12. Optionally, the polypeptide of the assays will comprise a domain of SEQ ID NOS:7 to 12. Either SEQ ID NOS:7 to 12 or a domain thereof can be  
25 covalently linked to a heterologous protein to create a chimeric protein used in the assays described herein.

Modulators of protein activity can be tested using polypeptides as described above, either recombinant  
30 or naturally occurring. The protein can be isolated, expressed in a cell, expressed in a membrane derived from a cell, expressed in tissue or in an animal, either recombinant or naturally occurring.

Changes in ion flux may be assessed by determining changes in polarization (i.e., electrical potential) of the cell or membrane expressing a protein of the invention. One means to determine changes in cellular polarization is by measuring changes in current (thereby measuring changes in polarization) with voltage-clamp and patch-clamp techniques, e.g., the "cell-attached" mode, the "inside-out" mode, and the "whole cell" mode (see, e.g., Ackerman et al., *New Engl. J. Med.* 336:1575-1595 (1997)). Whole cell currents are conveniently determined using the standard methodology (see, e.g., Hamil et al., *Pflugers. Archiv.* 391:85 (1981). Other known assays include: radiolabeled ion flux assays and fluorescence assays using voltage-sensitive dyes (see, e.g., Vestergaard-Bogind et al., *J. Membrane Biol.* 88:67-75 (1988); Gonzales & Tsien, *Chem. Biol.* 4:269-277 (1997); Daniel et al., *J. Pharmacol. Meth.* 25:185-193 (1991); Holevinsky et al., *J. Membrane Biology* 137:59-70 (1994)). Generally, the compounds to be tested are present in the range from 1 pM to 100 mM.

The practice of using a reporter gene to analyze nucleotide sequences that regulate transcription of a gene-of-interest is well documented. In particular, the promoters of the genes of the present invention, i.e., SEQ ID NOS:1-5, contain receptor responsive elements. These elements can confer a particular effect of the transcription factor nuclear hormone receptor of the invention and can be read by a reporter system, as described below.

The demonstrated utility of a reporter gene is in its ability to define domains of transcriptional regulatory elements of a gene-of-interest. Reporter

genes express proteins that serve as detectable labels indicating when the control elements regulating reporter gene expression are up or down-regulated in response to outside stimuli.

5

By way of example, two types of reporter gene assay are discussed below. The first is a scorable reporter gene, whose expression can be quantified, giving a proportional indication of the level of expression supported by the genetic construct comprising the reporter gene. The second example is a selectable reporter gene. When expressed, the selectable reporter gene allows the host cell harboring the reporter gene to survive under restrictive conditions that would otherwise kill (or retard the growth of) the host cell.

Scorable reporter genes are typically used when the relative activity of a genetic construct is sought, whereas selectable reporters are used when confirmation of the presence of the reporter expression construct within the cell is desired.

Firefly luciferase expression systems have become widely used for quantitative analysis of transcriptional modulation in living cells (see, e.g., Wood, K.V. (1998) *Promega Notes* 65:14). In particular, recombinant cells comprising this reporter construct enable libraries of small molecules to be rapidly screened for those affecting specific aspects of cellular physiology, such as receptor function or intracellular signal transduction. deWet et al. (1987) *Mol. Cell. Biol.* 7:725; Wood, K.V. (1991) In: *Bioluminescence and Chemiluminescence: Current Status*, eds. P. Stanley and L. Kricka, John Wiley and Sons, Chichester, 11.

The luciferase assay could be used to screen any of the potential reagents listed above. For example, by placing the luciferase gene under the control of the SCD promoter, reagents that bind to the SCD protein can trigger a feedback loop modulating expression of the luciferase gene. Similarly, by creating a fusion protein comprising the luciferase and SCD coding sequences, siRNAs, antisense sequences and ribozymes targeted against the SCD gene can be screened, as any reagent acting on the SCD transcript will necessarily disrupt expression of the luciferase enzyme encoded in the same transcript.

Modulators will manifest themselves by altering the amount of light emitted by the luciferase-catalyzed hydrolysis of ATP, with up-modulators increasing the amount of light emitted (they induce increased luciferase production) and down-modulators decreasing the amount of light emitted (by inhibiting luciferase production) in proportion to the degree of expressional modulation (at least within the linear range limits of the assay). Luciferase assay kits and other reporter gene constructs suitable for use in the present invention are well known in the art and commercially available, e.g., Invitrogen and Promega. See, e.g., *Steady-Glo™ Luciferase Assay Reagent Technical Manual Luciferase Assay Reagent Technical Manual #TM051*, Promega Corporation.

As an example of a reporter gene assay, the NurRE from the SCD target POMC can be inserted in front of the SV40 promoter in a pGL3-promoter construct (Promega) so that the luciferase can be induced by NR4A1 expression. Down-regulation of SCD can be assayed to

determine whether the SCD protein level correlates to its ability to induce reporter gene expression. For further description of reporter system assays, see, for example, Unit 9.6 of Current Protocols in Molecular Biology, (John Wiley & Sons, 1997).

A number of selectable marker systems can be used in the present invention, including but not limited to the herpes simplex virus thymidine kinase (Wigler, et al., 1977, *Cell* 11:223), hypoxanthine-guanine phosphoribosyltransferase (Szybalska & Szybalski, 1962, *Proc. Natl. Acad. Sci. USA* 48:2026), and adenine phosphoribosyltransferase (Lowy, et al., 1980, *Cell* 22:817) genes can be employed in tk<sup>-</sup>, hgp<sup>+</sup> or ap<sup>+</sup> cells, respectively. Also, antimetabolite resistance can be used as the basis of selection for dhfr, which confers resistance to methotrexate (Wigler, et al., 1980, *Natl. Acad. Sci. USA* 77:3567; O'Hare, et al., 1981, *Proc. Natl. Acad. Sci. USA* 78:1527); gpt, which confers resistance to mycophenolic acid (Mulligan & Berg, 1981, *Proc. Natl. Acad. Sci. USA* 78:2072); neo, which confers resistance to the aminoglycoside G-418 (Colberre-Garapin, et al., 1981, *J. Mol. Biol.* 150:1); and hyg<sup>+</sup>, which confers resistance to hygromycin (Santerre, et al., 1984, *Gene* 30:147) genes.

Typically, selectable markers are included in expression cassettes comprising the target gene or construct to be incorporated into the host cell. The selectable marker may be under the control of the same promoter as the target construct, e.g., as part of a fusion protein or polycistronic transcript; or may be under the control of an independent promoter.



As suggested above, the purpose of the selectable marker is to confer selectable growth characteristics on cells that are able to express it. By including the selectable marker in the same nucleic acid comprising the target gene or construct, the selectable marker will be included in any cell transformed with the target. Therefore, by selecting for the growth characteristics conferred by the selectable marker, cells transfected with the target can be selected.

10

Real-time PCR assays, as described in the examples herein, take advantage of those cycles of a normal PCR reaction where the DNA being amplified is increasing at a logarithmic rate and hence proportional to the amount of DNA present. Several kits are commercially available for performing real-time PCR. One such kit is the TaqMan® assay.

15

The TaqMan® assay takes advantage of the 5' nuclease activity of Taq DNA polymerase to digest a DNA probe annealed specifically to the accumulating amplification product. TaqMan® probes are labeled with a donor-acceptor dye pair that interacts via fluorescence energy transfer. Cleavage of the TaqMan® probe by the advancing polymerase during amplification dissociates the donor dye from the quenching acceptor dye, greatly increasing the donor fluorescence. All reagents necessary to detect two allelic variants can be assembled at the beginning of the reaction and the results are monitored in real time. In an alternative homogeneous hybridization based procedure, molecular beacons are used for allele discriminations. Molecular beacons are hairpin-shaped oligonucleotide probes that report the presence of specific nucleic acid molecules in

20

25

30

homogeneous solutions. When they bind to their targets they undergo a conformational reorganization that restores the fluorescence of an internally quenched fluorophore. (See, e.g., Heid, C.A., Stevens, J., Livak, K.J. and Williams, P.M. Real time quantitative PCR. *Genome Res.* 6:986-994 (1996); Gibson, U.E.M., Heid, C.A. and Williams, P.M. A novel method for real time quantitative RT-PCR. *Genome Res.* 6:995-1001 (1996)).

10               Northern blot methods allow RNA isolated from cells of interest to be separated using gel electrophoresis techniques. After separation, nucleic acids are transferred to membranes and hybridized with radio-labeled nucleotide probes. For analysis of  
15 expression maps, poly A (adenylyl) probed are used, which hybridize to mRNA species present on the blot.

                  The present invention includes both traditional and expression map Northern blotting. Expression of SEQ  
20 ID NOS:1 to 6 and other genes of interest can be tracked using probes specific for these genes. Expression mapping can be used to monitor alterations in gene expression in response to protein-specific binding agents.

25               Methods of RNA isolation are taught in, for example, Ausubel, F. M. et al., *Current Protocols in Molecular Biology*, Volume 1, pp. 4.1.1-4.2.9 and 4.5.1-4.5.3, John Wiley & Sons, Inc., 1993. Northern blot  
30 analysis is routine in the art and is taught in, for example, Ausubel, F. M. et al., *Current Protocols in Molecular Biology*, Volume 1, pp. 4.2.1-4.2.9, John Wiley & Sons, Inc., 1996.

Through the use of high density oligonucleotide arrays, expression profiles for individual cells can be rapidly obtained and compared. High density arrays are particularly useful for monitoring expression control at the transcriptional, RNA processing and degradation level. The fabrication and application of high density arrays in gene expression monitoring have been disclosed in, for example, WO 97/10365, WO 92/10588, U.S. Pat. No. 6,040,138 incorporated herein for all purposes by reference. In some embodiments using high density arrays, high density oligonucleotide arrays are synthesized using methods such as the Very Large Scale Immobilized Polymer Synthesis (VLSIPS) disclosed in U.S. Pat. No. 5,445,934. Each oligonucleotide occupies a known location on a substrate. A nucleic acid target sample is hybridized with a high density array of oligonucleotides and then the amount of target nucleic acids hybridized to each probe in the array is quantified. One preferred quantifying method is to use confocal microscope and fluorescent labels. The GeneChip.RTM. system (Affymetrix, Santa Clara, Calif.) is particularly suitable for quantifying the hybridization; however, it will be apparent to those of skill in the art that any similar systems or other effectively equivalent detection methods can also be used.

High density arrays are suitable for quantifying a small variations in expression levels of a gene in the presence of a large population of heterogeneous nucleic acids. Such high density arrays can be fabricated either by *de novo* synthesis on a substrate or by spotting or transporting nucleic acid sequences onto specific locations of substrate. Nucleic acids are purified and/or isolated from biological

materials, such as a bacterial plasmid containing a cloned segment of sequence of interest. Suitable nucleic acids are also produced by amplification of templates. As a nonlimiting illustration, polymerase chain reaction, and/or in vitro transcription, are suitable nucleic acid amplification methods.

Synthesized oligonucleotide arrays are particularly preferred for this invention.

Oligonucleotide arrays have numerous advantages, as opposed to other methods, such as efficiency of production, reduced intra- and inter array variability, increased information content and high signal-to-noise ratio.

15

An "antisense compound or molecule" refers to such compound or molecule that includes a polynucleotide that is complementary to a target sequence of choice and capable of specifically hybridizing with the target molecules. The term antisense includes a "ribozyme," which is a catalytic RNA molecule that cleaves a target RNA through ribonuclease activity. Antisense nucleic acids hybridize to a target polynucleotide and interfere with the transcription, processing, translation or other activity of the target polynucleotide. An antisense nucleic acid can inhibit DNA replication or DNA transcription by, for example, interfering with the attachment of DNA or RNA polymerase to the promoter by binding to a transcriptional initiation site or a template. It can interfere with processing of mRNA, poly(A) addition to mRNA or translation of mRNA by, for example, binding to regions of the RNA transcript such as the ribosome binding site. It can promote inhibitory mechanisms of the cells, such as promoting RNA

degradation via RNase action. The inhibitory polynucleotide can bind to the major groove of the duplex DNA to form a triple helical or "triplex" structure. Methods of inhibition using antisense polynucleotides  
5 therefore encompass a number of different approaches to altering expression of specific genes that operate by different mechanisms (see, e.g., Helene & Toulme, *Biochim. Biophys. Acta.*, 1049:99-125 (1990)).

10           The antisense compounds that may be used in connection with this embodiment of the present invention preferably comprise between about 8 to about 30 nucleobases (i.e., from about 8 to about 30 linked nucleosides), more preferably from about 12 to about 25  
15 nucleobases, and may be linear or circular in configuration. They may include oligonucleotides containing modified backbones or non-natural internucleoside linkages. Preferred modified oligonucleotide backbones include, for example,  
20 phosphorothioates, chiral phosphorothioates, phosphorodithioates, phosphotriesters, aminoalkylphosphotriesters, methyl and other alkyl phosphonates including 3'-alkylene phosphonates and chiral phosphonates, phosphinates, phosphoramidates  
25 including 3'-amino phosphoramidate and aminoalkylphosphoramidates, thionophosphoramidates, thionoalkylphosphonates, thionoalkylphosphotriesters, and boranophosphates having normal 3'-5' linkages, 2'-5' linked analogs of these, and those having inverted  
30 polarity wherein the adjacent pairs of nucleoside units are linked 3'-5' to 5'-3' or 2'-5' to 5'-2'. Methods of preparing antisense compounds are well known in the art (see, for example, U.S. Patent No. 6,210,892).

The present invention also provides a method of increasing inhibition of cancer cells, the method comprising introducing into the cell an effective amount of an expression vector comprising a sequence of  
5 nucleotides that encodes a ribozyme having a substrate binding sequence disclosed herein or an siRNA. The expression vector is preferably administered in combination with a suitable carrier. After the vector has been administered, the ribozyme or siRNA is expressed  
10 in the cell.

This method can be applied to a subject with cancer. Administration of the vector into the subject can be by any suitable route including oral, sublingual  
15 intravenous, subcutaneous, transcutaneous, intramuscular, intracutaneous, and the like. Any of a variety of non-toxic, pharmaceutically acceptable carriers can be used for formulation including, glucose, lactose, gum acacia, gelatin, mannitol, starch paste, magnesium trisilicate,  
20 talc, corn starch, keratin, colloidal silica, potato starch, urea, dextrans, and the like. The formulated material may take any of various forms such as injectable solutions, sterile aqueous or non-aqueous solutions, suspensions or emulsions, tablets, capsules, and the  
25 like.

As used herein, the phrase "effective amount" refers to a dose of the deliverable sufficient to provide circulating concentrations high enough to impart a  
30 beneficial effect on the recipient, which is an increase of inhibition of cancer cells.

The specific therapeutically effective dose level for any particular subject and deliverable depends

upon a variety of factors including the severity of the infection, the activity of the specific compound administered, the route of administration, the rate of clearance of the specific compound, the duration of treatment, the drugs used in combination or coincident with the specific compound, the age, body weight, sex, diet and general health of the patient, and like factors well known in the medical arts and sciences. Dosage levels typically range from about 0.001 up to 100 mg/kg/day; with levels in the range of about 0.05 up to 10 mg/kg/day.

The present invention also provides an antibody with binding specificity for a protein of the invention, such as SEQ ID NOS:7 to 12, or any molecule with 80%, 85%, 90% or 95% or more sequence identity with SEQ ID NOS:7 to 12, or any fragment of 10 or more contiguous amino acids of SEQ ID NOS:7 to 12. The antibody can have a binding specificity for a protein or peptide (i.e., amino acid sequence) encoded by SEQ ID NOS:1 to 6, or any molecule with 80%, 85%, 90% or 95% or more sequence identity with SEQ ID NOS:1 to 6.

As used herein, the term "antibody" refers to a polypeptide comprising a framework region from an immunoglobulin gene or fragments thereof that specifically binds and recognizes an antigen. The recognized immunoglobulin genes include the kappa, lambda, alpha, gamma, delta, epsilon, and mu constant region genes, as well as the myriad immunoglobulin variable region genes. Light chains are classified as either kappa or lambda. Heavy chains are classified as gamma, mu, alpha, delta, or epsilon, which in turn define

the immunoglobulin classes, IgG, IgM, IgA, IgD and IgE, respectively.

An exemplary immunoglobulin (antibody)  
5 structural unit comprises a tetramer. Each tetramer is composed of two identical pairs of polypeptide chains, each pair having one "light" (about 25 kDa) and one "heavy" chain (about 50-70 kDa). The N-terminus of each chain defines a variable region of about 100 to 110 or  
10 more amino acids primarily responsible for antigen recognition. The terms variable light chain ( $V_L$ ) and variable heavy chain ( $V_H$ ) refer to these light and heavy chains respectively.

15 Antibodies can exist as intact immunoglobulins or as a number of well-characterized fragments produced by digestion with various peptidases. Thus, for example, pepsin digests an antibody below the disulfide linkages in the hinge region to produce  $F(ab)'_2$ , a dimer of Fab  
20 which itself is a light chain joined to  $V_H-C_H1$  by a disulfide bond. The  $F(ab)'_2$  may be reduced under mild conditions to break the disulfide linkage in the hinge region, thereby converting the  $F(ab)'_2$  dimer into an Fab' monomer. The Fab' monomer is essentially Fab with part  
25 of the hinge region (see *Fundamental Immunology* (Paul ed., 3d ed. 1993)). While various antibody fragments are defined in terms of the digestion of an intact antibody, one of skill will appreciate that such fragments may be synthesized *de novo* either chemically or by using  
30 recombinant DNA methodology.

Thus, the term antibody, as used herein, also includes antibody fragments either produced by the modification of whole antibodies, or those synthesized *de*



5  
novo using recombinant DNA methodologies (e.g., single chain Fv) or those identified using phage display libraries (see, e.g., McCafferty et al., *Nature* 348:552-554 (1990)).

- 5  
Methods of producing polyclonal and monoclonal antibodies are known to those of skill in the art. See, e.g., Coligan (1991) *Current Protocols in Immunology* Wiley/Greene, NY; and Harlow and Lane (1989) *Antibodies: A Laboratory Manual*, Cold Spring Harbor Press, NY; Stites et al. (eds.) *Basic and Clinical Immunology* (4th ed.) Lange Medical Publications, Los Altos, CA; Huse et al. (1989) *Science* 246:1275-1281; and Ward et al. (1989) *Nature* 341:544-546, and references cited therein.
- 10  
Techniques for the production of single chain antibodies (U.S. Patent 4,946,778) can be adapted to produce antibodies to polypeptides of this invention.

- Also, transgenic mice, or other organisms such as other mammals, may be used to express humanized antibodies. Alternatively, phage display technology can be used to identify antibodies and heteromeric Fab fragments that specifically bind to selected antigens (see, e.g., McCafferty et al., *Nature* 348:552-554 (1990); Marks et al., *Biotechnology* 10:779-783 (1992)).
- 20  
25

- Methods of producing polyclonal and monoclonal antibodies that react specifically with a protein of the invention are known to those of skill in the art (see, e.g., Coligan, *Current Protocols in Immunology* (1991); Harlow & Lane, *supra*; Goding, *Monoclonal Antibodies: Principles and Practice* (2d ed. 1986); and Kohler & Milstein, *Nature* 256:495-497 (1975). Such techniques include antibody preparation by selection of antibodies
- 30

from libraries of recombinant antibodies in phage or similar vectors, as well as preparation of polyclonal and monoclonal antibodies by immunizing rabbits or mice (see, e.g., Huse et al., *Science* 246:1275-1281 (1989); Ward et al., *Nature* 341:544-546 (1989)).

A number of proteins comprising immunogens may be used to produce antibodies specifically reactive with SEQ ID NOS:6-10, or portions thereof. For example, recombinant NRA41 or an antigenic fragment thereof can be isolated, as is known in the art. Recombinant protein can be expressed in eukaryotic or prokaryotic cells as described above, and purified as generally described above. Recombinant protein is one embodiment of an immunogen for the production of monoclonal or polyclonal antibodies. Alternatively, a synthetic peptide derived from the sequences disclosed herein and conjugated to a carrier protein can be used as an immunogen. Naturally occurring protein may also be used either in pure or impure form. The product is then injected into an animal capable of producing antibodies. Either monoclonal or polyclonal antibodies may be generated, for subsequent use in immunoassays to measure the protein.

Methods of production of polyclonal antibodies are known to those of skill in the art. An inbred strain of mice (e.g., BALB/C mice) or rabbits is immunized with the protein using a standard adjuvant, such as Freund's adjuvant, and a standard immunization protocol. The animal's immune response to the immunogen preparation is monitored by taking test bleeds and determining the titer of reactivity to the protein of the invention. When appropriately high titers of antibody to the immunogen are obtained, blood is collected from the animal and

antisera are prepared. Further fractionation of the antisera to enrich for antibodies reactive to the protein can be done if desired (see Harlow & Lane, *supra*).

5            Monoclonal antibodies may be obtained by various techniques familiar to those skilled in the art. Briefly, spleen cells from an animal immunized with a desired antigen are immortalized, commonly by fusion with a myeloma cell (see Kohler & Milstein, *Eur. J. Immunol.*  
10 6:511-519 (1976)). Alternative methods of immortalization include transformation with Epstein Barr Virus, oncogenes, or retroviruses, or other methods well known in the art. Colonies arising from single immortalized cells are screened for production of  
15 antibodies of the desired specificity and affinity for the antigen, and yield of the monoclonal antibodies produced by such cells may be enhanced by various techniques, including injection into the peritoneal cavity of a vertebrate host. Alternatively, one may  
20 isolate DNA sequences which encode a monoclonal antibody or a binding fragment thereof by screening a DNA library from human B cells according to the general protocol outlined by Huse *et al.*, *Science* 246:1275-1281 (1989).

25            Monoclonal antibodies and polyclonal sera are collected and titered against the immunogen protein in an immunoassay, for example, a solid phase immunoassay with the immunogen immobilized on a solid support. Typically, polyclonal antisera with a titer of  $10^4$  or greater are  
30 selected and tested for their cross reactivity against proteins other than those of the invention or even other related proteins from other organisms, using a competitive binding immunoassay. Specific polyclonal antisera and monoclonal antibodies will usually bind with

a  $K_d$  of at least about 0.1 mM, more usually at least about 1  $\mu$ M, optionally at least about 0.1  $\mu$ M or better, and optionally 0.01  $\mu$ M or better.

5               Once specific antibodies are available, a protein of the invention can be detected by a variety of immunoassay methods. For a review of immunological and immunoassay procedures, see *Basic and Clinical Immunology* (Stites & Terr eds., 7th ed. 1991). Moreover, the  
10 immunoassays of the present invention can be performed in any of several configurations, which are reviewed extensively in *Enzyme Immunoassay* (Maggio, ed., 1980); and Harlow & Lane, *supra*.

15               Detection can be facilitated by coupling (i.e., physically linking) the antibody to a detectable substance. Examples of detectable substances include various enzymes, prosthetic groups, fluorescent materials, luminescent materials, bioluminescent  
20 materials, and radioactive materials. Examples of suitable enzymes include horseradish peroxidase, alkaline phosphatase, .beta.-galactosidase, or acetylcholinesterase; examples of suitable prosthetic group complexes include streptavidin/biotin and  
25 avidin/biotin; examples of suitable fluorescent materials include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride or phycoerythrin; an example of a luminescent material includes luminol; examples of  
30 bioluminescent materials include luciferase, luciferin, and aequorin, and examples of suitable radioactive material include  $^{125}$  I,  $^{131}$  I,  $^{35}$  S or  $^3$ H.

The antibodies are also useful for inhibiting receptor function, for example, blocking ligand binding. These uses can also be applied in a therapeutic context in which treatment involves inhibiting receptor function.

5 An antibody can be used, for example, to block ligand binding. Antibodies can be prepared against specific fragments containing sites required for function or against intact receptor associated with a cell.

10 As used herein, the phrase "binding specificity," in relationship to an antibody that binds to a protein or peptide, refers to a binding reaction which is determinative of the presence of the protein in the presence of a heterogeneous population of proteins

15 and other biologics. Thus, under designated immunoassay conditions, the specified antibody binds to a particular protein and does not bind significantly to other proteins present in the sample.

20 The following examples are intended to illustrate but not limit the present invention.

#### Example 1

25 This example shows that knockdown of SCD (SEQ ID NO:1; GenBank Accession #: NM\_005063) results in increased cancer cell death (apoptosis).

##### 1. General.

30 Stearoyl-CoA desaturase (SCD), also known as delta-9-desaturase, is an iron-containing enzyme that catalyzes the rate limiting step in the production of unsaturated fatty acids. The enzyme belongs to the

family of oxidoreductases. Its primary product is oleic acid, which is formed by desaturation of stearic acid. The ratio of stearic acid to oleic acid has been implicated in the regulation of cell growth and differentiation.

The SCD gene is located on chromosome 10, specifically 10q23-10q24, and spans approximately 24kb. It contains 6 exons and produces two different sized transcripts due to alternative poly A signals. The size of the mRNAs are 3.9 and 5.2 kb and they code for similar sized proteins. See Zhang et al., Biochem J., 340:255-264 (1999); Samuel et al., J. Biol. Chem., 277:59-66 (2002). SCD is highly expressed in several cancerous tissues, notably breast cancer (CGAP database).

## 2. Construction of siRNAs targeting SCD.

Three siRNA constructs were designed against SCD, using the Silencer(TM) siRNA Construction Kit manufactured by Ambion, Inc. (Austin, Texas). These siRNAs were designated CoA-A, CoA-B, and CoA-C, respectively, as described in Table 1 below.

Table 1

SiRNA	Binding position to SEQ ID NO:1	Sense sequence of siRNA
CoA-A	323-343	5' AAGGAGGAGATAAGTTGGAG 3' (SEQ ID NO:17)
CoA-B	1028-1048	5' AACAGTGCTGCCACCTCTTC 3' (SEQ ID NO:18)
CoA-C	1347-1367	5' AACCAGCCAGGCAGAGGTTTT 3' (SEQ ID NO:19)

### 3. siRNA knockdown of SCD in DLD-1 cells

The siRNAs described above were quantitated by absorbance at 260nm and diluted to a working concentration of 1 $\mu$ M. Each of the siRNAs, and a randomized scrambled control were transfected individually into DLD-1 cells seeded in 12 well plates at approximately 30% confluence. Transfection was done at a concentration of 10nM using Oligofectamine™. Forty eight hours after transfection, the cells were induced to undergo apoptosis by the addition of Fas agonistic antibody. Twenty four hours after induction, the cells were harvested and assayed for apoptosis by the DNA fragmentation ELISA assay.

The results of three independent experiments as described above are summarized in Figure 15. These show that knockdown of SCD by two of the three siRNAs (CoA-A and CoA-B -- SEQ ID NOS:17 and 18) induced apoptosis as compared to the scrambled control, with the CoA-A siRNA inducing the greatest level.

To confirm the correlation between the observed increase in apoptosis and knockdown of SCD, realtime PCR was performed to quantify SCD mRNA knockdown. RNA samples were isolated from the DLD-1 cells transfected with CoA-A (SEQ ID NO:17), CoA-B (SEQ ID NO:18) and the scrambled siRNA. Real time PCR was performed on the samples for 40 cycles and the values were normalized to 18S rRNA. The results from duplicate real time PCR experiments are shown in Figure 16. As can be seen, the amount of SCD mRNA in the DLD-1 cells was significantly reduced as a result of transfection with the CoA-A siRNA (SEQ ID NO:17).

#### **4. siRNA knockdown of SCD in A2058 cells (Fas-induced apoptosis)**

5                   To further assess the correlation between  
knockdown of SCD and an increase in cancer cell death,  
the same three siRNA constructs described above were then  
transfected into A2058 (melanoma) cells, which are known  
to be resistant to Fas-induced apoptosis. The protocol  
10 was the same as described above for the DLD-1 cells. As  
can be seen from Figure 17, knockdown of SCD by the CoA-A  
siRNA (SEQ ID NO:17) resulted in a high level of Fas  
induced apoptosis.

15                   To confirm these results, real-time PCR was  
performed, as described above in connection with DLD-1  
cells. RNA samples were isolated from A2058 cells  
transfected with CoA-A (SEQ ID NO:17), CoA-B (SEQ ID  
NO:18) and the scrambled siRNA. As can be seen from  
20 Figure 18, transfection with the CoA-A siRNA resulted in  
a significant reduction in SCD mRNA.

#### **5. siRNA knockdown of SCD in A2058 cells (TRAIL-induced apoptosis)**

25                   Further experiments were conducted to ascertain  
whether knockdown of SCD by the CoA-A siRNA (SEQ ID  
NO:17) would result in an increase in TRAIL-induced  
apoptosis .

30                   CoA-A (SEQ ID NO:17) and control siRNAs were  
transfected into A2058 cells, as described above. TRAIL  
was added forty eight hours after transfection. Twenty



four hours after the addition of the TRAIL, cells were harvested and assayed. In one experiment, apoptosis induction was measured by a DNA fragmentation ELISA assay. See Figure 19. In another experiment, apoptosis induction was assayed by means of Caspase 3/7 activity. See Figure 20. Both experiments showed that knockdown of SCD by siRNA results in an increase in TRAIL-induced apoptosis in A2058 cells, which are normally resistant to TRAIL-induced apoptosis.

## 6. Knockdown of SCD in non-cancerous cells

This experiment shows that knockdown of SCD does not result in an increase in either Fas-induced or TRAIL-induced apoptosis in non-cancerous cells. The three siRNAs described above (SEQ ID NOS:17 to 19) were transfected, as described above, into non-cancerous HuVec (human vascular endothelial) cells, and the cells were then tested for activation of both Fas and TRAIL induced apoptosis. The results are shown in Figure 21. As can be seen, no increase in either Fas-induced or TRAIL-induced apoptosis was observed. The readings never got as high as even 50% above the background (determined by mixing the substrate reagent with buffer in an ELISA well). From this data, it can be inferred that knockdown of SCD resulting in an increase in apoptosis induction is specific to cells that are cancerous.

## Example 2

This example shows that knockdown of CA12 (SEQ ID NO:2; GenBank Accession #: NM\_001218) results in increased cancer cell death (apoptosis).

## 1. General.

Carbonic anhydrase XII is a type-1 membrane protein. It is an enzyme with an extracellular catalytic domain which shows about 30 to 42% similarity to other known human carbonic anhydrases. However, unlike the catalytic domains of some of the other membrane bound carbonic anhydrases, the catalytic domain of carbonic anhydrase XII is truly catalytic and does not bind ligands. The gene is located on chromosome 15, specifically 15q22, and the size of the mRNA is 4.5 kb. See Ulmasov et al., PNAS, 97:14212-217 (2000).

## 2. Construction of siRNAs targeting CA12.

Five siRNA constructs were designed against CA12, using the Silencer(TM) siRNA Construction Kit manufactured by Ambion, Inc. (Austin, Texas). These siRNAs were designated CA12-1, CA12-2, CA12-3, CA12-4, and CA12-5, respectively, as shown below in Table 2 below.

Table 2

SiRNA	Binding position to SEQ ID NO:2	Sense sequence of siRNA
CA12-1	345-365	5' AAGGCTACAATCTGTCTGCCA 3' (SEQ ID NO:20)
CA12-2	731-751	5' AACATTGAAGAGCTGCTTCCG 3' (SEQ ID NO:21)
CA12-3	1285-1305	5' AATCTGCAAACATGCCAGGCC 3' (SEQ ID NO:22)
CA12-4	563-583	5' AACTCAGACCTTTATCCTGAC 3' (SEQ ID NO:23)
CA12-5	950-970	5' AAGTTCGATGAGAGGCTGGTA 3' (SEQ ID NO:24)

### 3. siRNA knockdown of CA12 in DLD-1 cells.

5           The siRNAs described above were quantitated by  
absorbance at 260nm and diluted to a working  
concentration of 1 $\mu$ M. Each of the siRNAs, and a  
randomized scrambled control, were transfected  
individually into DLD-1 cells seeded in 12 well plates at  
10 approximately 30% confluence. Transfection was done at a  
concentration of 10nM using Oligofectamine™. Forty eight  
hours after transfection, the cells were induced to  
undergo apoptosis by the addition of Fas agonistic  
antibody. Twenty four hours after induction, the cells  
15 were harvested and assayed for apoptosis by the DNA  
fragmentation ELISA assay.

          The results of three independent experiments as  
described above are summarized in Figure 23. These show  
20 that knockdown of SCD by two of the siRNAs, CA12-3 and  
CA12-5 (SEQ ID NOS: 22 and 24) resulted in a significant  
increase in apoptosis as compared to the scrambled  
control.

25           To confirm the correlation between the observed  
increase in apoptosis and knockdown of CA12, real-time  
PCR was performed to quantify CA12 mRNA knockdown. RNA  
samples were isolated from the DLD-1 cells transfected  
with CA12-3, CA12-4, CA12-5 (SEQ ID NOS:22 to 24) and the  
30 scrambled siRNA, as described above. Real-time PCR was  
performed on the samples for 40 cycles and the values  
were normalized to 18S rRNA. The results from duplicate  
real time PCR experiments are shown in Figure 24. As can  
be seen, the amount of CA12 mRNA in the DLD-1 cells was

significantly reduced as a result of transfection with the CA12-3 and CA12-5 siRNAs (SEQ ID NOS: 22 and 24).

#### 4. siRNA knockdown of CA12 in AsPC-1 cells.

5

To further assess the correlation between knockdown of CA12 and an increase in cancer cell death, the same five siRNA constructs described above were then transfected into AsPC-1 (pancreatic cancer) cells, which are known to be resistant to Fas-induced apoptosis. The protocol was the same as described above for the DLD-1 cells. The results of three independent transfection experiments are shown in Figure 25. As can be seen, knockdown of CA12 by the siRNAs resulted in an increase in Fas-induced apoptosis, with the CA12-3 and CA12-5 siRNAs (SEQ ID NOS:22 and 24) producing the highest level.

As was done in connection with the previous experiments, real-time PCR was performed to quantify CA12 mRNA knockdown. RNA samples were isolated from AsPC-1 cells transfected with the CA12-3, CA12-4, CA12-5 (SEQ ID NOS:22 to 24) and the scrambled siRNA, as described above. Real-time PCR was performed for 40 cycles and the values were normalized to 18S rRNA. The results of duplicate experiments are summarized in Figure 26. As can be seen, a reduction of CA12 mRNA occurred as a result of transfection with all three siRNAs, with the greatest reduction resulting from transfection with CA12-5.

### Example 3

This example shows that knockdown of PIK3R4 (SEQ ID NO:3; GenBank accession #: NM\_014602) inhibits the survival of cancer cells.

5

#### 1. General

PIK3R4 (phosphoinositide-3-kinase, regulatory subunit 4, p 150) encodes a predicted  
10 polypeptide of 1,358 amino acids (SEQ ID NO:8). The amino acid sequence is 29.6% identical to that of yeast vps15 protein. Northern blot analysis reveals that the gene is ubiquitously expressed as a 5.5-kb transcript. See Panaretou et al., J. Biol. Chem., 272:2477-2485  
15 (1997). Molecular epidemiology from the Virtual Northern database indicates that PIK3R4 is over-expressed in cancers including glioblastoma, colon, and brain cancers.

#### 2. Construction of siRNAs targeting PIK3R4

20

Four siRNA constructs targeting different regions of PIK3R4 mRNA were designed and transcribed in vitro, in a manner similar to that described in connection with the siRNAs of Examples 1 and 2. The  
25 sense sequences of these siRNAs are listed in Table 3 below.

Table 3

SirNA	Sequence of sirNA
A	AAGCTCTCTTCAACTCTCCTCCCTGTCTC (SEQ ID NO:25)
B	AATGGTCAACGTCCTCAAGGCCCTGTCTC (SEQ ID NO:26)
C	AACCTCAGCATCACTAAGACCCCTGTCTC (SEQ ID NO:27)
D	AACTGGAAGGATGACAGTGACCCTGTCTC (SEQ ID NO:28)

### 3. siRNA knockdown of PI3KR4 in A172 cells.

The four siRNAs targeting PI3KR4 and a scrambled control siRNA were in vitro transcribed and transfected into A172 (glioblastoma) cells, in 96 well plates (at a concentration of 15nM) For each individual siRNA, 6 wells were transfected, along with the control scrambled siRNA. Two days after transfection, the cells were treated with 10 nM C<sub>2</sub> ceramide (to induce apoptosis). 18 hours after the addition of C<sub>2</sub> ceramide, cell viability/survival was assayed by means of a WST-1 assay (Roche Molecular Biochemicals, Mannheim, Germany). The results (measuring decrease in cell survival) are shown in Figure 27. As can be seen, knockdown of PI3KR4 by siRNA D (SEQ ID NO:28) resulted in a statistically significant reduction in cell survival.

#### Example 4

This example shows that knockdown of PLD3 (SEQ ID NO:4; GenBank Accession #: NM\_012268) inhibits the survival of cancer cells.

##### 1. General

The phospholipase D family of proteins hydrolyzes phosphatidylcholine into phosphatidic acid (PA) and choline. PA acts as a second messenger in signal transduction and is further metabolized to diacylglycerol (DAG). Thus, the regulation of PLD could have profound effects on signal transduction. The locus for PLD3 maps to 19q13.2. Currently, it appears that the gene encodes

an mRNA of 2.1 Kb. The protein contains 437aa and has an approximate molecular weight of 43,000 daltons.

The expression of the phospholipase D3 (PLD3; SEQ ID NO:4) is greater in the brain than in non-nervous tissue and appears to be predominantly of neuronal origin. PLD3 expression is pronounced in mature neurons of the forebrain and appears to be turned on at late stages of neurogenesis as revealed by *in situ* hybridization analysis of PLD3 expression during postnatal development of the hippocampal formation and the primary somatosensory cortex. See Kim et al., Exp Mol Med., 35:38-45 (2003). Molecular epidemiology from the Virtual Northern database indicates that PLD3 is over-expressed in cancers including colon, kidney, stomach, and lung cancers.

## **2. Construction of siRNAs targeting PLD3**

Four siRNA constructs targeting different regions of PLD3 mRNA were designed and transcribed *in vitro*, in a manner similar to that described in connection with the siRNAs of Examples 1, 2 and 3 above. The sense sequences of these siRNAs are listed in Table 4 below.

30

Table 4

SiRNA	Binding position to SEQ ID NO:4	Sense sequence of siRNA
A	573-593	AAGCAGTGCTGGTGGAAAGCA (SEQ ID NO:29)
B	980-1000	AACATGGACTGGCGTTCCTG (SEQ ID NO:30)
C	1253-1273	AAGGCTCTACTCAACGTGGTG (SEQ ID NO:31)
D	1552-1572	AATCCCATATGCCCGTGTCAA (SEQ ID NO:32)

### 3. siRNA knockdown of PLD3 in A172 cells.

5                   The four siRNAs targeting PLD3 and a scrambled control siRNA were in vitro transcribed and transfected into A172 (glioblastoma) cells, in 96 well plates (at a concentration of 15nM), one day after seeding. For each individual siRNA, 6 wells were transfected, along with

10 the control scrambled siRNA. Two days after transfection, the cells were treated with 10 nM C<sub>2</sub> ceramide (to induce apoptosis. 18 hours after the addition of C<sub>2</sub> ceramide, cell viability/survival was assayed by means of the WST-1 assay (Roche Molecular

15 Biochemicals, Mannheim, Germany). The results (measuring decrease in cell survival) are shown in Figure 29A. As can be seen, knockdown of PLD3 by all four siRNAs resulted in a statistically significant reduction in cell survival.

20

To confirm the correlation between the observed decrease in cell survival and knockdown of PLD3, real-time PCR was performed to quantify PLD3 mRNA knockdown. Sub-confluent A172 cells were transfected with the

25 control siRNA and three siRNAs against PLD3 using Oligofectamine™. Cells were harvested 48 hrs after



transfection. mRNA was then made and subject to real-time RT-PCR analysis. The results, as seen in Figure 29B (message level normalized to 18S) indicate that a significant reduction in PLD3 mRNA occurred as a result of the transfection with siRNAs targeting PLD3.

### Example 5

This example shows that knockdown of HSPD1 (SEQ ID NO:5; GenBank Accession #: NM\_002156) inhibits proliferation of cancer cells.

#### 1. General.

HSPD1 (SEQ ID NO:5), also referred to as chaperonin 60, GROEL, HSP60 (heat shock protein 60) and SPG13, is a highly conserved gene located at chromosome 2q33.1. HSPD1 encodes an ATPase of 573 amino acids. It belongs to a family of chaperon proteins which contain a TCP-1 (T-complex protein)/cpn60 chaperonin domain. HSPD1 is an abundant protein found primarily in the mitochondria, although 15% to 20% is found in the cytosol.

Heat shock proteins (Hsps) are a set of highly conserved proteins, some of which are constitutively expressed in mammalian cells (e.g. Hsp90 and Hsp60), whereas others are induced in response to stress (e.g. Hsp72 and Hsp27). The inducible expression, or constitutive over-expression of the latter group of Hsps is known to promote cell survival through inhibition of apoptosis. Samali, and Orrenius, Cell Stress, Chap. 3, 228-236 (1998). The constitutively expressed Hsps act as

chaperones for other cellular proteins, binding to nascent polypeptides to prevent premature folding, maintain enzymes in active conformation, and facilitate translocation of proteins into organelles. Lindquist and  
5 Craig, Annu. Rev. Genet., 22:631-677 (1988). It has been shown that reduction in HSPD1 precipitates translocation of bax to the mitochondria facilitating apoptosis. Gupta et al., Circulation, 106:2727-33 (2002). The functions of Hsp60 are strongly dependent upon Hsp10,  
10 which binds to the Hsp60 molecule to regulate its substrate binding and ATPase activity. Ryan et al., Int. Rev. Cytol., 174:127-193 (1997). Molecular epidemiological data indicate that HSPD1 is over-expressed in many cancerous tissues including brain,  
15 ovary, breast, prostate, stomach, muscle, and placenta.

## 2. Expression of HSPD1 in cancer cell lines

20 Real-time PCR (RT-PCR) experiments were performed in order to assess the expression levels of HSPD1 in several experimental cell lines: Hela (cervical cancer); AsPC-1 (pancreatic cancer); DLD-1 (colon cancer), PC3 (prostate cancer); T47D (breast cancer),  
25 A2058 (melanoma) and U87 (glioma). RNA samples from sub-confluent cultures were prepared by standard procedures, and RT-PCR assays were performed using these RNAs. The results are shown in Figure 31 (in Figure 31, HSPD1 mRNA level is normalized to the internal control ribosomal 18S  
30 RNA level; relative mRNA levels are presented using arbitrary units).

### 3. Construction of siRNAs targeting HSPD1

Four *in vitro* transcribed siRNA constructs (designated A,B,C,D) against HSPD1 were generated using the Silencer(TM) siRNA Construction Kit manufactured by Ambion, Inc. (Austin, Texas). The sequences of these constructs are outlined below in Table 5:

Table 5

SiRNA	Sense sequence of siRNA
A	5'- AACACAAATGAAGAAGCTGGG -3' (SEQ ID NO:33)
B	5'- AAATTGCCAATGCTCACCCTA -3' (SEQ ID NO:34)
C	5'- AAGAACAGTGATTATTGAGCA -3' (SEQ ID NO:35)
D	5'- AATTGCTGAACTTAAAAAGCA -3' (SEQ ID NO:36)

### 4. Knockdown of HSPD1 in A2058 cells

The effect of HSPD1 knockdown by the siRNAs described above was first tested in A2058 cells. Briefly, 5e3 A2058 cells were seeded in 96-well plates and transfected on the following day, using 0.4 $\mu$ l Oligofectamine™, with the four siRNAs and a control siRNA. 24 hours post-transfection, media was changed to 1% serum. 48 hours post-transfection, cells were treated with either 1% serum or 1% serum + brefeldin A (0.05 $\mu$ g/ml) and incubated for 3 days (Brefeldin A induces sensitization to apoptosis through the endoplasmic reticulum (ER) pathway). Cell survival was then measured using the WST-1 assay (Roche Molecular Biochemicals, Mannheim, Germany). The results of three experiments and nine data points are shown in Figure 32A (treatment only with 1% serum) and Figure 32B (treatment with a combination of 1% serum and brefeldin A).

As can be seen, knockdown of HSPD1 by three of the four siRNAs (A, B and C; SEQ ID NOS:33 to 35) caused a decrease in cell viability when assayed in 1% serum, and knockdown of HSPD1 by all four siRNAs caused a significant decrease in cell viability when assayed in 1% serum and brefeldin A.

#### 5. Knockdown of HSPD1 in AsPC-1 cells

The experiments described above were repeated using the pancreatic cancer cell line, AsPC-1. Briefly, 5e3 AsPC1 cells were seeded in 96-well plates and transfected with 10nM in vitro transcribed siRNAs on the following day, using 0.4µl Oligofectamine™. 24 hours post-transfection, media was changed to 1% serum. 48 hours post-transfection, cells were treated with either 1% serum or 1% serum + brefeldin A (0.05µg/ml) and incubated for 3 days. Cell viability was then measured using the WST-1 assay. The results of three experiments and nine data points are shown in Figure 33A (1% serum only) and Figure 33B (1% serum and brefeldin A). In this case, knockdown of HSPD1 by two of the siRNAs (A and B; SEQ ID NOS:33 and 34) resulted in a statistically significant decrease in cell viability both in 1% serum and in the combination of 1% serum and brefeldin A.

#### 6. Knockdown of HSPD1 in PC3 cells

The same experiments described above were repeated using the prostate cancer cell line, PC3. The results of three experiments and nine data points are shown in Figure 34. As can be see, in this cell line, as well, knockdown of HSPD1 by two siRNAs (A and B; SEQ ID

NOS:33 and 34) resulted in a decrease in cell viability as measured by the WST-1 assay.

#### 7. Knockdown of HSPD1 in A2058 cells (TRAIL-induced apoptosis)

An additional experiment was conducted to test the effect of knockdown of HSPD1 on the ability of the cells to undergo apoptosis. In this experiment, A2058 cells were used. Briefly, 5e3 cells were seeded in 96-well plates and transfected on the following day with 10nM in vitro transcribed siRNAs using 0.2 $\mu$ l Oligofectamine™. The siRNAs used were A and B (SEQ ID NOS: 35 and 36) and a control siRNA. 48 hours post-transfection, killer TRAIL (10ng/ml) was added to the cells for 5 hours. Caspase 3/7 activity was then measured using the Apo-ONE assay (Promega). The results from one experiment and 3 data points are shown in Figure 35. As can be seen, knockdown of HSPD1 by one of the siRNAs (A; SEQ ID NO:33) resulted in a significant increase in apoptosis.

#### Example 6

This example shows that knockdown of ZPK (SEQ ID NO:4; GenBank Accession #: NM\_006301) inhibits the proliferation of cancer cells.

##### 1. General.

ZPK (SEQ ID NO:6), also named MAP3K12 (mitogen-activated protein kinase kinase kinase 12), DLK, and MUK, is a Ser/Thr and Tyr protein kinase. It is one of the

leucin-zipper MAP kinase (leucine zipper bearing kinase - LZK), and belongs to the mixed-lineage kinases (MLKs) family which has been shown to involve the JNK activation pathway. ZPK interacts with the GTP bound form of RAC and CDC42, regulators (activators) of JNK pathway. Xu et al., Mol Cell Biol., 21:4713 (2001). ZPK has also been found to modulate mitogen-activation protein kinase signaling to allow or prevent differentiation. Ruiz-Hidalgo et al., Exp. Cell Res., 274:1788 (2002); Hebert et al., J Biol Chem., 275:32482 (2000). Nihalani et al., J. Biol. Chem., 275:7273 (2000).

The gene for ZPK is located at chromosome 12q13.13, and encodes a protein of 859AA in length (SEQ ID NO:12). It has two transcript variants, variant 1 (AK094195; SEQ ID NO:40; Figure 41) and variant 2 (NM\_006301; SEQ ID NO:6). These two variants are identical for most of the sequence with the exception of a 99 base pair insertion in the coding region for variant 1, resulting in a 33 amino acids insertion (see Figure 37; and Figure 42, which shows the amino acid sequence, SEQ ID NO:41, coded by variant 1 of ZPK (SEQ ID NO:40)). All of the claimed methods that apply to variant 2 (SEQ ID NO:6) and its encoded amino acid sequence (SEQ ID NO:12) and segments, as disclosed herein, apply as well to variant 2 (SEQ ID NO:40) and its encoded amino acid sequence (SEQ ID NO:41).

The resulting Molecular epidemiological data indicate that ZPK is overexpressed in many cancers including brain, bone, prostate, colon, thyroid, ovary, uterus, and pancreatic cancers.

## 2. Expression of ZPK in cancer cell lines.

Real time PCR (RT-PCR) experiments were performed in order to assess the expression levels of ZPK in several experimental cell lines: Hela (cervical cancer); AsPC-1 (pancreatic cancer); MCF7 (breast cancer); DLD-1 (colon cancer), A2058 (melanoma), T47D (breast cancer); JHMG (glioma). RNA samples from sub-confluent cultures were prepared by standard procedures and RT-PCR was performed using these RNAs. The results (Figure 38) show that ZPK is expressed in all cell lines tested (Figure 38). (In Figure 38, the ZPK message level is normalized to the internal control ribosomal 18S RNA level; relative mRNA levels are presented using arbitrary units. Ct (GAPDH) = 17, Ct (ZPK) = 24 in Hela.)

## 3. Construction of siRNAs targeting ZPK

Three siRNAs against ZPK were generated by in vitro transcription, using the Silencer(TM) siRNA Construction Kit manufactured by Ambion, Inc. (Austin, Texas). Their sequences are described in Table 6 below:

Table 6

SiRNA	Sense sequence of siRNA
A	5'- GCAGCAGCAGGAAGACCUtt -3' (SEQ ID NO:37)
B	5'- GAGGAACUGGUGAUGAGGAtt -3' (SEQ ID NO:38)
C	5'- GUCAGAAACGUGGCAUCUCtt -3' (SEQ ID NO:39)

#### 4. siRNA knockdown of ZPK in HeLa cells

The three siRNA against ZPK described above and a control siRNA were transiently transfected into HeLa cells, and the transfected cells were trypsinized and plated into 96-well plates 24 hrs post transfection (D0). Cell growth in liquid culture was quantified on day 7 by alamarBlue™ staining (BioSource International, Inc., Camarillo, CA). The results of two independent experiment and six data points are summarized in Figure 39A (P-values for all three siRNAs <0.01). These experiments show that knockdown of ZPK by siRNA results in a significant reduction in liquid culture (anchorage dependent) growth of HeLa cancer cells.

15

The same number of transfected cells from the experiment described above was also plated into 96-well plates for and assayed for soft agar (anchorage independent) growth. Following a week of soft-agar culture, cells were stained with alamarBlue™. The results of two independent experiments and six data points are summarized in Figure 39B. These show that knockdown of ZPK by siRNA also results in a significant reduction in anchorage independent growth of HeLa cells.

25

To confirm the observed results of the previous two experiments, sub-confluent HeLa cells were transfected with a control siRNA and the three siRNAs against ZPK using Oligofectamine™. Cells were harvested 48 hours after transfection. mRNA was then made and subject to real-time RT-PCR (Taqman®) analysis. The results as shown in Figure 39C (message level normalized to 18S) indicate a significant reduction in ZPK mRNA by the siRNAs targeting ZPK as compared to the control.

30



## 5. siRNA knockdown of ZPK in DLD-1, AsPC-1 and PC3 cells

The effect of ZPK knockdown on anchorage independent growth was also tested in several additional  
5 cell lines. The three siRNAs against ZPK described above  
and a control siRNA were transfected into DLD-1, AsPC-1,  
and PC3 cells in 96-well plate using Oligofectamine™.  
The cells were harvested 24 hrs later, and 1/10 of cells  
per well were seeded in 96 well plates for soft-agar  
10 cultures. Following a week of soft-agar culture, the  
cells were stained with alamarBlue™. The results of two  
independent experiments and six data points for DLD-1 and  
PC3 cells are shown in Figures 40A and 40B, respectively;  
the results of a single experiment and three data points  
15 for AsPC-1 cells are shown in Figure 40C. As can be seen  
from these figures, knockdown of ZPK in all three cell  
lines resulted in a statistically significant reduction  
in anchorage-independent growth.

20 All references made herein, including articles,  
patent applications and any other publications, are  
incorporated by reference in their entirety.

25

30

We claim:

1. A method of identifying a molecule that inhibits cancer cells, comprising:
  - a. introducing said molecule into cells, wherein said molecule binds to a compound comprising SEQ ID NOS:1, 2, 3, 4, 5 or 6, or comprising the RNA correlate of SEQ ID NOS:1, 2, 3, 4, 5, 6 or 40; and
  - b. measuring the level of inhibition of said cells, where an increase in level indicates said molecule inhibits cancer cells.
2. The method of claim 1, wherein said compound comprises SEQ ID NOS:1, 2, 3, 4, 5, 6 or 40.
3. The method of claim 1, wherein said compound comprises the RNA correlate of SEQ ID NOS:1, 2, 3, 4, 5, 6 or 40.
4. The method of claim 1, wherein said compound consists of SEQ ID NOS:1, 2, 3, 4, 5, 6 or 40.
5. The method of claim 1, wherein said compound consists of the RNA correlate of SEQ ID NOS:1, 2, 3, 4, 5, 6 or 40.
6. The method of claim 1, wherein said inhibition is measured by apoptosis, where an increase in the level of apoptosis indicates that said molecule inhibits said cells.
7. The method of claim 1, wherein said inhibition is measured by proliferation, where a decrease in the rate

of cell division indicates that said molecule inhibits said cells.

8. The method of claim 1, wherein said cells are  
5 cancerous.

9. The method of claim 1, wherein said cells are selected from the group consisting of HeLa cells, A2058 cells, DLD1 cells, T47D cells, ASPC1 cells and JH-MG cells.  
10

10. A method of identifying a molecule that inhibits cancer cells, comprising:

a. introducing said molecule into cells, wherein said  
15 molecule binds to a compound comprising an amino acid sequence selected from the group consisting of residues 96 to 321 of SEQ ID NO:7; residues 32 to 289 of SEQ ID NO:8; residues 26 to 320 of SEQ ID NO:9; residues 143 to 170 of SEQ ID NO:10; residues 358 to 384 of SEQ ID NO:10;  
20 residues 47 to 550 of SEQ ID NO:11; residues 158 to 405 of SEQ ID NO:12; and residues 175 to 414 of SEQ ID NO:41; and

b. measuring the level of inhibition of said cells, where  
25 an increase in level indicates said molecule inhibits cancer cells.

11. The method of claim 10, wherein said molecule binds to a compound selected from the group consisting of  
30 residues 96 to 321 of SEQ ID NO:7; residues 32 to 289 of SEQ ID NO:8; residues 26 to 320 of SEQ ID NO:9; residues 143 to 170 of SEQ ID NO:10; residues 358 to 384 of SEQ ID NO:10; residues 47 to 550 of SEQ ID NO:11; residues 158

to 405 of SEQ ID NO:12; and residues 175 to 414 of SEQ ID NO:41.

12. The method of claim 10, wherein said molecule binds  
5 to a compound comprising a sequence selected from the group consisting of SEQ ID NOS:7, 8, 9, 10, 11, 12 and 41.

13. The method of claim 10, wherein said molecule binds  
10 to a compound selected from the group consisting of SEQ ID NOS:7, 8, 9, 10, 12 and 41.

14. The method of claim 10 wherein said inhibition is measured by apoptosis, where an increase in the level of  
15 apoptosis indicates that said molecule inhibits said cells.

15. The method of claim 10, wherein said inhibition is measured by proliferation, where a decrease in the rate  
20 of cell division indicates that said molecule inhibits said cells.

16. A method of inhibiting cancer cells, comprising introducing into said cells a molecule that binds to a  
25 compound comprising SEQ ID NOS:1, 2, 3, 4, 5, 6 or 40, or comprising the RNA correlate of SEQ ID NOS:1, 2, 3, 4, 5, 6 or 40, whereby the level of cell inhibition is increased.

30 17. The method of claim 16, wherein said compound comprises SEQ ID NOS:1, 2, 3, 4, 5, 6 or 40.

18. The method of claim 16, wherein said compound comprises the RNA correlate of SEQ ID NOS:1, 2, 3, 4, 5, 6 or 40.
- 5 19. The method of claim 16, wherein said molecule is siRNA.
20. The method of claim 19, wherein said siRNA is selected from the group consisting of SEQ ID NOS:17 to  
10 39.
21. A method of identifying a molecule that inhibits cancer cells, comprising:
- 15 a. introducing said molecule into cells, wherein said molecule down-modulates a compound comprising the RNA correlate of SEQ ID NOS:1, 2, 3, 4, 5, 6 or 40; and
- b. measuring the level of down-modulation of said  
20 compound, where an increase in level of down-modulation indicates that said molecule inhibits cancer cells.
22. The method of claim 21, wherein said down-modulation is measured by a reporter assay using a reporter gene.  
25
23. The method of claim 22, wherein said reporter gene expresses a protein selected from the group consisting of  $\beta$ -lactamase, luciferase, green fluorescent protein,  $\beta$ -galactosidase, secreted alkaline phosphatase, human  
30 growth hormone and chlororamphenicol acetyltransferase.
24. A method of identifying a molecule that inhibits cancer cells, comprising:

- a. introducing said molecule into cells, wherein said molecule down-modulates a compound comprising SEQ ID NOS:7, 8, 9, 10, 11, 12 or 41; and
- 5 b. measuring the level of down-modulation of said compound, where an increase in level of down-modulation indicates that said molecule inhibits cancer cells.
25. The method of claim 24, wherein said molecule binds  
10 to a compound selected from the group consisting of SEQ ID NOS:7, 8, 9, 10, 11, 12 and 41.
26. The method of claim 24, wherein said down-modulation is measured by an immunoassay using an antibody specific  
15 to said compound.
27. The method of claim 26, wherein said immunoassay is an immunofluorescence, immunochemistry or immunoprecipitation assay.

20

# **ABSTRACT**

The present invention is directed to methods of identifying agents useful for inhibiting cancer cells.

5 The agents bind to one of the proteins described herein, or to a gene or mRNA encoding such protein. The invention is also directed to methods for inhibiting cancer cells by administering an agent that binds to one of the proteins described herein, or to a gene or mRNA  
10 encoding such protein.

# Hairpin Ribozyme

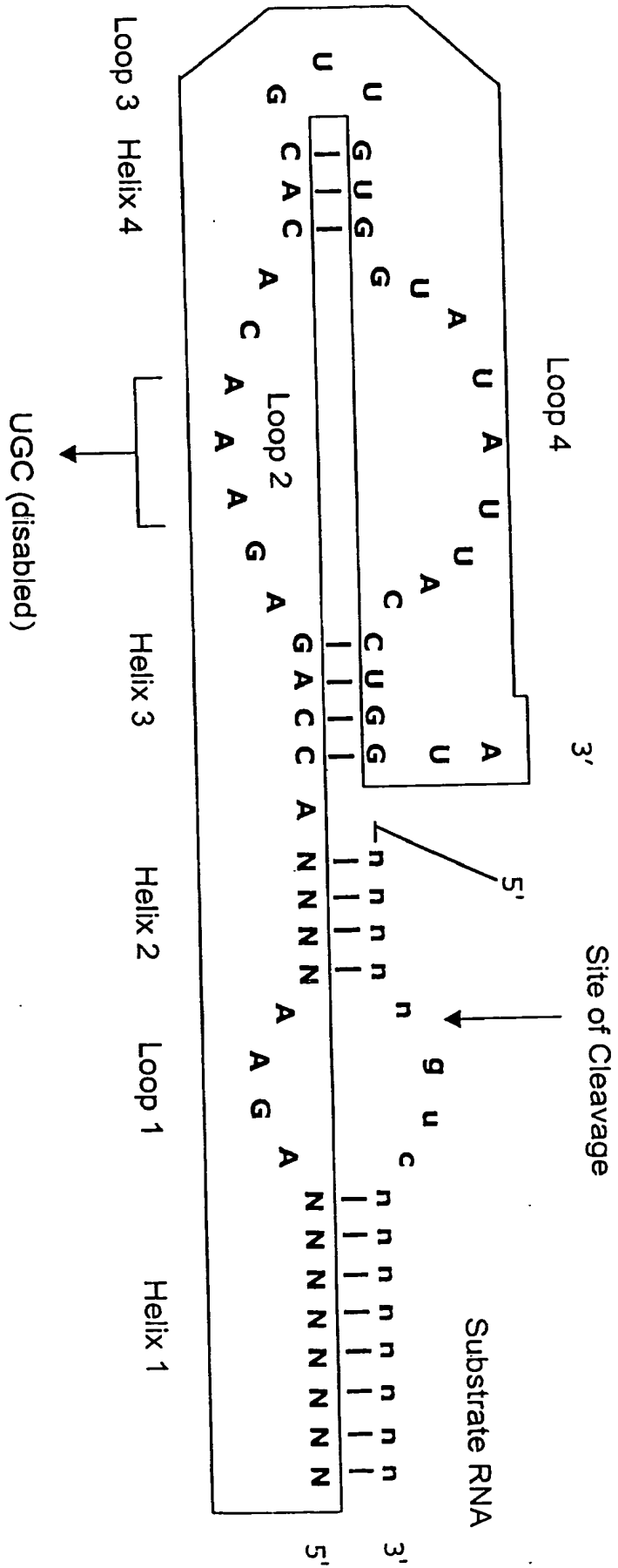


Figure 1



# Hammerhead Ribozyme

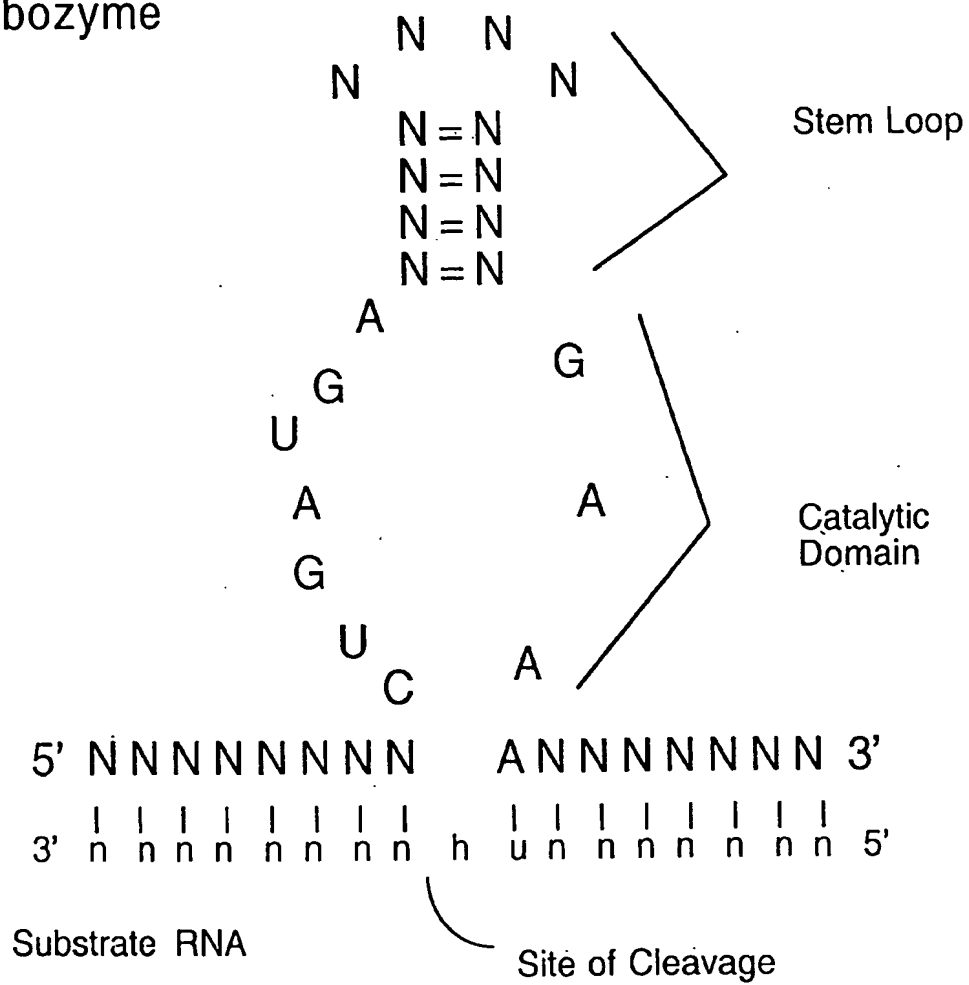


Figure 2

SCD (DNA sequence; SEQ ID NO:1)

```

1  gtggtgtcgg tgtcggcagc atccccggcg ccttgcgtcg gtcgccggag cctcggcct
61  ctgtttctct cccctcccgc ccttacctc cagcgggac cggccgcgc agtcaactcc
121  tcgcactttg cccctgcttg gcagcggata aaagggggct gaggaatac cggacacgtc
181  caccggttgc cagctctagc ctttaaattc ccggctcggg acctccacgc accgggctag
241  cgccgacaac cagctagcgt gcaaggcgcc ggggtcagc gcgtaccggc gggcttcgaa
301  accgcagtc tccggcgacc ccgaactccg ctccggagcc tcagccccct ggaaagtgat
361  cccggcatcg gagagccaag atgccggccc acttgctgca ggacgatata tctagctcct
421  ataccaccac caccacatt acagcgctc cctccagggt cctgcagaat ggaggagata
481  agttggagac gatccccctc tacttggaag acgacattcg cctgatata aaagatgata
541  tatatgaccc cactacaag gataagggaag gcccaagccc caagggtgaa tatgtctgga
601  gaaacatcat ccttatgtct ctgctacact tgggagccct gtatgggac actttgattc
661  ctacctgcaa gttctacacc tggctttggg gggatttcta ctattttgtc agtgccctgg
721  gcataacagc aggagctcat cgtctgtgga gccaccgctc ttacaaagct cggctgcccc
781  tacggctctt tctgatcatt gccaacacaa tggcattcca gaatgatgtc tatgaatggg
841  ctctgaccca ccgtgcccac cacaagtttt cagaaacaca tgctgatcct cataattccc
901  gacgtggctt tttcttctct cactgtgggt ggctgcttgt gcgcaaacac ccagctgtca
961  aagagaaggg gactacgcta gacttgctcg acctagaagc tgagaaactg gtgatgttcc
1021  agaggaggta ctacaaacct ggcttgcgtc tgatgtgctt catcctgccc acgcttgtgc
1081  cctgggtattt ctgggtgaa acttttcaaa acagtgtgtt cgttgccact ttcttgcgat
1141  atgctgtggt gcttaatgcc acctggctgg tgaacagtgc tgcccacctc ttcggatata
1201  gtccttatga caagaacatt agccccggg agaatacctt ggtttcactt ggagctgtgg
1261  gtgagggtct ccacaactac caccactcct ttcctatga ctactctgcc agtgagtacc
1321  gctggcacat caacttcacc acattcttca ttgattgcat ggccgcccctc ggtctggcct
1381  atgaccggaa gaaagtctcc aaggccgcca tcttggccag gattaaaaga accggagatg
1441  gaaactacaa gagtggctga gtttggggtc cctcaggttc ctttttcaaa aaccagccag
1501  gcagagggtt taatgtctgt ttattaacta ctgaataatg ctaccaggat gctaaagatg
1561  atgatgttaa cccattccag tacagtattc ttttaaaatt caaaagtatt gaaagccaac
1621  aactctgcct ttatgatgct aagctgatat ttttcttct cttatcctct ctctcttcta
1681  ggcccattgt cctccttttc actttaatcg cctcctttc ccttattgcc tcccaggcaa
1741  gcagctggtc agtctttgct cagtgtccag cttccaaagc ctagacaacc tttctgtagc
1801  ctaaaacgaa tggcttttgc tccagataac tctctttcct tgagctgttg tgagctttga
1861  agtaggtggc ttgagctaga gataaaacag aatcttctgg gtagtccctt gttgattatc
1921  ttcagcccag gcttttgcta gatggaatgg aaaagcaact tcatttgaca caaagcttct
1981  aaagcnaggt aaattgtcgg gggagagagt tagcatgtat gaatgtaagg atgagggag
2041  cgaaggaacc tctcgccatg atcagacata cagctgccta cctaagagg acttcaagcc
2101  ccaccacata gcatgcttcc tttctctcct ggctcggggg aaaaagtggc tgcggtgttt
2161  ggcaatgcta attcaatgcc gcaacatata gttgaggccg aggataaaga aaagacattt
2221  taagtttgta gtaaaagtgg tctctgctgg ggaaggggtt tcttttctt ttttctttaa
2281  taacaaggag atttcttagt tcatatatca agaagtcttg aagttgggtg tttccagaat
2341  tggtaaaaac agcagctcat agaattttga gtattccatg agctgctcat tacagttcct
2401  tctcttttct gctctgccat cttcaggata ttggttcttc ccctcatagt aataagatgg
2461  ctgtggcatt tccaaacatc caaaaaaagg gaaggattta aggaggtgaa gtcgggtcaa
2521  aaataaaata tatatacata tatacattgc ttagaacggt aaactattag agtatttccc
2581  ttccaaagag ggatgttttg aaaaaactct gaaggagagg aggaattagt tgggatgcca
2641  atttctcttc cactgctgga catgagatgg agaggctgag ggacaggatc tataggcagc
2701  ttctaagagc gaacttcaca taggaaggga tctgagaaca cgttcagggg ttgagaagggt
2761  tactgagtga gttattggga gtcttaataa actagatatt aggtccattc attaataggt
2821  tccagtttct ccttgaaatg agtaaaaact agaaggcttc tctccacagt gttgtgcccc
2881  ttcactcatt tttttttgag gagaaggggg tctctgttaa catctagcct aaagtataca
2941  aactgcctgg ggggcagggt taggaatctc ttcactacct tgattcttga ttctgtgctc
3001  taccctgtct gtcccttttc tttgaccaga tctttctctt ccctgaacgt tttcttctt
3061  ccctggacag gcagcctcct ttgtgtgtat tcagaggcag tgatgacttg ctgtccaggc
3121  agtccccctc tgcacacaga atgctcaggg tcaactgaac actgcttctc ttttgaaagt
3181  agagctagct gccactttca cgtggcctcc gcagtgtctc cacctacacc cctgtgctcc
3241  cctgccacac tgatggctca agacaaggct ggcaaaccct cccagaaaca tctctggccc

```

FIGURE 3

```

3301 agaaagcctc tctctccctc cctctctcat gagaagccaa gcgctcatgt tgagccagtg
3361 ggccagccac agagcaaaag aggggtttatt ttcagtcccc tctctctggg tcagaaccag
3421 agggcatgct gaatgcccc tgcttacttg gtgaggggtgc cccgcctgag tcagtgtctt
3481 cagctggcag tgcaatgctt gtagaagtag gaggaaacag ttctcactgg gaagaagcaa
3541 gggcaagaac ccaagtgcct cacctcgaaa ggaggccctg ttccctggag tcagggtgaa
3601 ctgcaaagct ttggctgaga cctgggattt gagataccac aaaccctgct gaacacagtg
3661 tctgttcagc aaactaacca gcattcccta cagcctaggg cagacaatag tatagaagtc
3721 tggaaaaaaa caaaaacaga atttgagaac cttggaccac tcctgtccct gtagtctagt
3781 catcaaagca gaagtctggc tttgctctat taagattgga aatgtacact accaaacact
3841 cagtccactg ttgagcccca gtgctggaag ggaggaaggc ctttcttctg tgttaattgc
3901 gtagaggcta caggggttag cctggactaa aggcattcctt gtctttgagc tattcacctc
3961 agtagaaaag gatctaaggg aagatcactg tagtttagtt ctggtgacct gtgcacctac
4021 ccttggtgaa tgtctgtctg tatttctaata tccacaggtc atcagatgcc tgcttgataa
4081 tatataaaca ataaaaacaa ctttctacttc ttcctattgt aatcgtgtgc catggatctg
4141 atctgtacca tgaccctaca taaggctgga tggcacctca ggctgagggc cccaatgtat
4201 gtgtggctgt ggtgtgtgggt gggagtgtgt ctgctgagta aggaacacga ttttcaagat
4261 tctaaagctc aattcaagtg acacattaat gataaactca gatctgatca agagtccgga
4321 tttctaacag tccttgcttt ggggggtgtg cctgcaactt agctcagggt cttacatct
4381 tttctaatac cagtgttgca tatgagcctg ccctcactcc ctctgcagaa tccctttgca
4441 cctgagacct tactgaagtg gctggtagaa aaaggggcct gagtggagga ttatcagtat
4501 cacgatttgc aggtattcct tctgggcttc attctgaaa cttttgttag ggctgctttt
4561 cttaaagtgc cacatttgat ggagggtgga aataatttga atgtatttga tttataagtt
4621 tttttttttt tttgggttaa aagatggttg tagcatttaa aatggaaaat tttctccttg
4681 gtttgctagt atcttgggtg tattctctgt aagtgtagct caaataggtc atcatgaaag
4741 gttaaaaaag cgagggtggc atgttatgct ggtggttgcc agggcctcca accactgtgc
4801 cactgacttg ctgtgtgacc ctgggcaagt cacttaacta taagggtgcct cagttttcct
4861 tctgttaaaa tggggataat aatactgacc tacctcaaag ggcagttttg aggcattgact
4921 aatgcttttt agaaagcatt ttgggatcct tcagcacagg aattctcaag acctgagtat
4981 tttttataat aggaatgtcc accatgaact tgatacgctc gtgtgtccca gatgtgtca
5041 ttagtctata tggttctcca agaaactgaa tgaatccatt ggagaagcgg tggataacta
5101 gccagacaaa atttgagaat acataaacia cgcattgcca cggaaacata cagaggatgc
5161 cttttctgtg attgggtggg attttttccc tttttatgtg ggatatagta gttacttgtg
5221 acaagaataa ttttggata atttctatta atatcaactc tgaagctaata tgtactaatc
5281 tgagattgtg tttgttcata ataaaagtga agtgaatctg attgactg

```

FIGURE 3 (cont.)

CA12 (DNA sequence; SEQ ID NO:2)

```

1 gtactcgcca cggcaccag gctgcgcga cgggtcccc gtgtgcagct ggagagcgag
61 cggccaccgg gagccccgg cacagccgc gcccgcccc caggagccc cgaagatgcc
121 ccggcgcagc ctgcacgcg cggcctgtct cctgctggtg atcttaaagg aacagccttc
181 cagcccgccc ccagtgaacg gttccaagt gacttatttt ggtcctgatg gggagaatag
241 ctggtccaag aagtaccctg cgtgtgggg cctgctgcag tccccatag acctgcacag
301 tgacatcctc cagtatgac ccagcctcac gccctcgcag ttccaaggct acaatctgtc
361 tgccaacaag cagtttctcc tgaccaaaa tggccattca gtgaagctga acctgccctc
421 ggacatgcac atccaggggc tccagtctcg ctacagtgcc acgcagctgc acctgcactg
481 ggggaaccgg aatgaccgc acggctctga gcacaccgtc agcggacagc acttcgcgcg
541 cgagctgcac attgtccatt ataactcaga cctttatcct gacgccagca ctgccagcaa
601 caagtacaga ggccctcgtg tcctggctgt tctcattgag atgggctcct tcaatccgtc
661 ctatgacaag atcttcagtc accttcaaca tgtaaagtac aaaggccagg aagcattcgt
721 cccgggattc aacattgaag agctgcttcc ggagaggacc gctgaatatt accgctaccg
781 ggggtccctg accacacccc cttgcaaccc cactgtgctc tggacagttt tccgaaaccc
841 cgtgcaaat tcccaggagc agctgctggc tttggagaca gccctgtact gcacacacat
901 ggacgaccct tccccagag aaatgatcaa caacttccgg caggtccaga agttcgatga
961 gaggtggtg tacacctcct tctcccaagt gcaagtctgt actgcggcag gactgagtct
1021 gggcatcate ctctcactgg ccctggctgg cattcttggc atctgtattg tgggtggtggt
1081 gtccatttgg cttttcagaa ggaagagtat caaaaaagg gataacaagg gagtcattta
1141 caagccagcc accaagatgg agactgagcc ccacgcttga ggtccccgga gctccccggc
1201 acatccagga aggaccttgc tttggaccct acacacttgc gctctctgga cacttgcgac
1261 acctcaagggt gttctctgta gctcaatctg caaacatgcc aggcctcagg gatcctctgc
1321 tgggtgcctc cttgccttgg gaccatggcc accccagagc catccgatcg atggatggga
1381 tgcactctca gaccaagcag caggaattca aagctgcttg ctgtaactgt gtgagattgt
1441 gaagtgtct gaattctgga atcacaaacc aagccatgct ggtgggcat taatggttgg
1501 aaaacacttt catccggggc tttgccagag cgtgctttca agtgccttgg aaagtctgct
1561 gcttctccaa gctttcagac aagaatgtgc actctctgct taggttttgc ttgggaaact
1621 caacttcttt cctctggaga cggggcatct ccctctgatt tccttctgct atgacaaaac
1681 ctttaatctg caccttacia ctcggggaca aatggggaca ggaaggatca agttgtagag
1741 agaaaaaaga aaacaagaga tataattgt gatataatag ggacactttc acagtcctgt
1801 cctctggatc acagacactg cacagacctt agggaaatggc aggttcaagt tccacttctt
1861 ggtggggatg agaagggaga gagagctaga gggacaaaga gaatgagaag acatggatga
1921 tctggggagag tctcactttg gaatcagaat tggaaacaca ttctgtttat caagccataa
1981 tgtaaggaca gaataataca atattaagtc caaatccaac ctctgtcag tggagcagtt
2041 atgttttata ctctacagat tttacaaata atgaggtgtg tccttgaaaa tgtgtgtgtg
2101 ctgtgtcctg gaggagacat gagttccgag atgacccaat ctgcctttga atctggagga
2161 aataggcaga aacaaaatga ctgtagaact tattctctgt aggccaaatt tcatttcagc
2221 cacttctgca ggatccctac tgccaacctg gaatggagac ttttatctac ttctctctct
2281 ctgaagatgt caaatcgtgg tttagatcaa atataattca agctataaaa gcaggaggtt
2341 atctgtgcag ggggctggca tcatgtattt aggggcaagt aataatggaa tgctactaag
2401 atactccata ttcttccccg aatcacacag acagtttctg acaggcgcaa ctctccatt
2461 ttcttccgc aggtgagaac cctgtggaga tgagtcagtg ccatgactga gaaggaaccg
2521 acccctagtt gagagcacct tgcagttccc cgagaacttt ctgattcaca gtctcatttt
2581 gacagcatga aatgtcctct tgaagcatag ctttttaaat atctttttcc ttctactcct
2641 ccctctgact ctaagaattc tctctctggt aatcgcttga acccaggagg cggaggttgc
2701 agtaagccaa ggtcatgcca ctgcactcta gcctgggtga cagagcgaga ctccatctca
2761 aaaaaaaaaa aaaaa

```

FIGURE 4

PIK3R4 (DNA sequence; SEQ ID NO:3)

```

1 gcacgagggg agttcggcgt ttgctggggc tgcagcagct gaagtgtagt gttttcttgg
61 gactggcggt ctgcacttct ctcccgggtt ccatctcccc ccgcccgggtg gtgaggecct
121 cgaggagggc tcggacgggt gtagcgatcc gcgctagagg aagacgaggc ccgggaacgc
181 atgtccccc aaggcaggtta gggggctgga ggggtcaa atcccgggtac ttgtggagac
241 tcttttagcgt ggcttcttct ctctgctgag accccgagag ctttcccagt tctcctccca
301 ggaccaccgg ggctcctgaa gatcgggact tttctgcgcc cctccaccaa cagcccatct
361 cctgtctatg aagaaagacc ctctgtagaa acaacttccc cgctgctgac gcgttttccc
421 gtcccgcccc cgaagtatgc tactatgacc tcgttgtgag cctctgaacg attttgacac
481 tttcccgagg cctagggtat tatatcctaa ccttactaaa gaccacagag gtgcttgcca
541 ttatgggaaa tcagcttgct ggcattgctc cctcccagat cttttctgta gagagttatt
601 tttcagatat tcatgacttt gaatatgata aaagcctggg gactactcgg ttttttaaag
661 ttgctcgagc caagcaccga gaaggcctgg tcgttgtgaa ggtttttgca attcaggatc
721 ccacattgcc tttaaccagc tataaacaag agctggagga actgaaaatc aggcttaatt
781 ctgcacagaa ttgtctacct ttccagaaag catcagaaaa agcatctgag aaagcagcta
841 tgctcttttag gcagtatgtg cgagacaatc tctatgatcg catcagtacc cgtccattct
901 tgaataacat tgagaagcgc tggattgctt tccagatcct gacagctgtg gaccaagcac
961 acaaactctg agttcgtcat ggggacatca agactgagaa tgtgatggtc accagttgga
1021 attgggttct tctaactgat tttgccagtt ttaagcccac ttatcttcca gaagacaacc
1081 cggcagattt caattatttc tttgacacat cacggaggag aacttgctat attgctcctg
1141 aacgttttgt tgatgggtgg atgtttgcca ctgagttaga atatatgaga gatccttcaa
1201 ctccgcttgt agacttaaat agcaatcaga gaacaagagg agagttgaag agagcaatgg
1261 acatcttttc agcaggttgt gtgatagctg agctttttac agaaggtgta ccattatttg
1321 atctctctca acttttggtt tatagaaatg gacatttttt ccctgaacaa gtgctaaata
1381 aaattgaaga tcacagtatc agagaattgg taactcagat gattcacctg gagccagata
1441 aacgttttaga ggcagaagat tacttaaaac agcagcgtgg caatgccttt cctgaaatat
1501 tttacacttt tcttcagccc tacatggccc agtttgccaa ggaaacgttt ctttctgcag
1561 atgagcgtat tctggttata cggaaggatt tgggcaacat tattcacaat ctctgtggac
1621 atgatctgcc agaaaaagcc gaaggagagc ctaaggaaaa tgggctggtt atcttggtat
1681 ctgttataac atcctgccta cagaccctta aatactgtga ttccaaacta gctgcttttg
1741 aactgattct tcatttggtt ccaagattaa gtgttgaaat ctttttggtt cgtattactc
1801 catatctttt catattcagc aatgactctg ttccatagggt gagggctgaa gccttgagga
1861 cgttgaccaa agttcttgct ctctgctaaa aggttctctg taatgatata aatatttatc
1921 cggaatacat tctgccaggc atagcccact tagcccaaga tgatgctact atcggtgac
1981 tagcctatgc tgaaaacata gctctgctgg cagaaacagc tctgagattc ctggaattag
2041 tacagttaaa aaatcttaat atggaaaatg accccaataa tgaagaaata gatgaggtta
2101 cacatccaaa tggaaattat gacacagagc tccaagcctt acatgaaatg gtccagcaga
2161 aagttgttac tttgctaagt gaccctgaaa atattgtaaa acaaaccttg atggaaaatg
2221 gaataacacg gctgtgtgta ttctttggac gtcagaaagc caacgatgtt ttgttgtccc
2281 acatgattac tttcctaaat gataagaatg attggcatct acgtggagca ttttttgata
2341 gtatagttgg tgttgctgcc tatgttggtt ggcaaagctc ctcaattctc aagcctctgc
2401 tgcaacaagg tcttagtgat gctgaggaat ttgtcattgt gaaagctctt tatgccctta
2461 cttgtatgtg ccagttagga ctgctacaaa aaccccatgt ttacgaattt gccagtgata
2521 ttgccccctt cctgtgtcat cccaatttat ggatacgtta tggtgccgtg ggatttatca
2581 cagtggtagc tcgtcaaata agtacagctg atgtctactg taaactgatg ccttatcttg
2641 accatataat taccacacca ataatacaga ttgaaagaaa acttgttctg ctcaagtgtt
2701 taaaggaacc agtaagtcgt tctatatattg attatgcttt gaggtctaaa gatattacta
2761 gcttggtcag acatcttcac atgcgtcaga agaaacgaaa tggttctctt cccgactgcc
2821 ctccgccaga ggatcctgcc atagcacagc ttctgaagaa gttgctctca caggaatga
2881 cagaggaaga ggaagacaaa cttctggcac tgaaagactt catgatgaaa tctaataaag
2941 caaaggccaa tatagtggac cagagccatc ttcatgatag tagtcagaaa ggtgtaattg
3001 acttggcagc tttaggcata actgggagac aagttgatct tgtaaaacc aaacaagaac
3061 cagatgacaa acgggcagaa aaacatgtaa aacaagactc aaatgtaaat gaagaatgga
3121 aaagcatgtt tgggtcactg gaccaccaa acatgccaca ggccctacct aaagggagtg
3181 atcaggaggt gattcagact gggaaacctc ctcgttccga gtcctctgct ggcatttgtg

```

FIGURE 5

3241 tccctttgtc aacttcttca caggttccag aagtgacaac tgtccaaaat aaaaaaccag  
3301 taataccggt tttaagtagt acaatcttac catccaccta tcagattcga attacaactt  
3361 gtaaaactga acttcagcaa ctcatccagc aaaagcggga gcagtgcaat gctgagagaa  
3421 tagctaagca gatgatggaa aatgctgaat gggagagtaa accaccacca cctggatggc  
3481 gtcctaaagg gctgttagtt gcccatcttc atgagcataa atctgctgtg aatcgaatta  
3541 gagtctctga tgaacactca ctttttgcaa catgttcaaa tgatggcaca gtgaaaatct  
3601 ggaacagtca aaagatggag gggaagacca ccactaccag atctattctt acatacagcc  
3661 gaattggagg acgagtcaag acgctcacat tctgccagg ctccactat ttagccatag  
3721 catctgataa tgggtctgtc cagcttcttg gaattgaggc ttctaagctg cccaagtctc  
3781 ctaaaatcca tcctctacaa agcagaattc tagatcagaa ggaggacggt tgtgttgtgg  
3841 atatgcatca cttcaactct ggagcacagt ctgttcttgc ctatgccact gtgaatggct  
3901 ctctgggttg ctgggacctt aggtcttcaa gcaatgcgtg gactttaaag catgatttaa  
3961 agtcgggcct catcacttcc tttgctgtgg acatccacca atgctggctc tgcattggta  
4021 caagcagtgg taccatggct tgttgggaca tgaggttcca gttgccatt tcaagtcact  
4081 gtcaccttcc cagggtctga atcagacgcc tctcaatgca ccctctgtat cagtcctggg  
4141 tgattgcagc tgttcagggc aacaacgaag tgtccatgtg ggacatggag actggtgaca  
4201 gaagatttac tctctgggcc agcagtgcac caccacttcc tgaattacag ccttctctc  
4261 atagcgtcca tgggtatctac tgtagtcctg cagatggaaa tcctatccta ctaacagctg  
4321 gctcagatat gaaaataagg ttttgggact tggcttacc agaaaggctc tatgttgttg  
4381 caggaagtac tagttcccca tctgtgtcct actacaggaa aataattgaa ggcaactgaag  
4441 ttgtccagga aattcagaat aagcagaaag taggaccaag tgatgacacc cctcgaaggg  
4501 gccagagtc cctgcccggt ggacatcatg acatcatcac tgatgtcgcc acattccaga  
4561 ccacacaggg ctcatcgta actgcttcta gagatgggat tgtgaagggtg tggaaaataaa  
4621 acctactgat ttgtataaat tttaatagtt ataaatataa tactataact cgagaaaagg  
4681 catcttctaga gaacagattc atttgtctaa ttttcaaaat tatgtctcca tattactgtt  
4741 tcatgactga ctgactaaat gacacccaaa atgggttaaga tgtacttgac tagtttactt  
4801 atgcatctct ttgcaagaat cagccagcca acaatgtctg ggatttttat tgtatatgtt  
4861 atagaggtga gaaatgtaaa atatgaaaat gaatatgttt attttgtatt gaaaaagatg  
4921 gttgaaaaga tgggttgaag ctattatagt ataaacacat ttttgctatt aaaaatgcta  
4981 ttcaaagcag ttaaactgta aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa  
5041 tcgagggggg gcccggtacc

FIGURE 5 (cont.)

PLD3 (DNA sequence; SEQ ID NO:4)

```

1   ctctttataa tttagtttcc atagaagtta tatgtgcatt taaaaaaatt caatgctgga
61  gcgaccgtgt ctggggagcc gagccccgct tctcgctgcg gtgagcccgg actggggcac
121 gcaactgcga gactccccgc tgcagtgggc ggagtcccac aggccccgcc ctcctccca
181 ccctcgttca gcctgtccag acagaagtgc gggcccagcg gaggtagcag cagacgcctg
241 agagcgaggg cgaggccctc aggggtttgga gaccctgaca caccacctt ctcacctggg
301 ctctgcgtat ccccagcct tgagggaaga tgaagcctaa actgatgtac caggagctga
361 aggtgcctgc agaggagccc gccaatgagc tgcccatgaa tgagattgag gcgtggaagg
421 ctgcggaaaa gaaagcccgc tgggtcctgc tggtcctcat tctggcgggt gtgggcttcg
481 gagcctgatg actcagctgt ttctatggga atacggcgac ttgcatctct ttgggccccaa
541 ccagcgccca gccccctgct atgacccttg cgaagcagt ctggtggaaa gcattcctga
601 gggcctggac ttccccaatg cctccacggg gaacccttcc accagccagg cctggctggg
661 cctgctcgcc ggtgcgcaca gcagcctgga catcgctcc ttctactgga ccctaccaa
721 caatgacacc cacacgcagg agccctctgc ccagcagggt gaggagggtc tccggcagct
781 gcagaccctg gcaccaaagg gcgtgaacgt ccgcatcgct gtgagcaagc ccagcgggcc
841 ccagccacag gcggacctgc aggtctctgt gcagagcggg gccaggtcc gcattggtga
901 catgcagaag ctgacctatg gcgtcctgca taccaagttc tgggtggtgg accagaccca
961 cttctacctg ggcagtgcc aatggactg gcgttctact acccaggtca aggagctggg
1021 cgtggtcatg tacaactgca gctgcctggc tcgagacctg accaagatct ttgaggccta
1081 ctggttccct ggccaggcag gcagctccat cccatcaact tggccccggg tctatgacac
1141 cgcgtacaac caagagacac caatggagat ctgcctcaat ggaacccctg ctctggccta
1201 cctggcgagt gcgccccac ccctgtgtcc aagtggcgc actccagacc tgaaggctct
1261 actcaacgtg gtggacaatg cccggagttt catctacgtc gctgtcatga actacctgcc
1321 cactctggag ttctccacc ctcacagggt ctggcctgcc attgacgatg ggctgcggcg
1381 ggccacctac gagcgtggcg tcaaggtgcg cctgtctatc agctgtctgg gacactcgga
1441 gccatccatg cgggccttcc tgctctctct ggctgcccct cgtgacaacc ataccactc
1501 tgacatccag gtgaaactct ttgtggtccc cgcggatgag gccaggtctc gaatcccata
1561 tgcccggtgc aaccacaaca agtacatggt gactgaacgc gccacctaca tcggaacctc
1621 caactggtct ggcaactact tcacggagac ggcgggcacc tcgctgctgg tgacgcagaa
1681 tgggaggggg ggcctgcgga gccagctgga ggccattttc ctgagggaact gggactcccc
1741 ttacattcat gaccttgaca cctcagctga cagcgtgggc aacgcctgcc gcctgctctg
1801 agggccgatc cagtgggcag gccaaaggct gctgggcccc cgcggacca ggtgctctgg
1861 gtcacgggtc ctgtccccgc acccccgtt ctgtctgccc cattgtggct cctcaggctc
1921 tctcccctgc tctccacct ctacctccac ccccaccggc ctgacgctgt ggccccggga
1981 cccagcagag ctgggggagg gatcagcccc caaagaaatg ggggtgcatg ctggcctgcc
2041 ccctggccca ccccacttt ccagggcaaa aagggccag ggttataata agtaaataac
2101 ttgtctgtaa aaaaaaaaaa aaaaaaaaaa a

```

FIGURE 6

HSPD1 (DNA sequence; SEQ ID NO:5)

```

1   ggcacgagggc gacgacctgt ctgccgagc gcacgccttg ccgccgcccc gcagaaatgc
61  ttcggttacc cacagtcttt cgccagatga gaccggtgtc caggggtactg gctcctcatc
121 tccactcgggc ttatgcaaaa gatgtaaaat ttggtgcaga tgcccagacc ttaatgcttc
181 aagggtgtaga ccttttagcc gatgctgtgg ccggtacaat gggggccaaag ggaagaacag
241 tgattattga gcagagttgg ggaagtccca aagtaacaaa agatgggtgtg actggtgcaa
301 agtcaattga cttaaaagat aaatacaaaa acattggagc taaacttgtt caagatgttg
361 ccaataacac aaatgaagaa gctggggatg gcactaccac tgctactgta ctggcacgct
421 ctatagccaa ggaaggcttc gagaagatta gcaaagggtc taatccagtg gaaatcagga
481 gaggtgtgat gttagctgtt gatgctgtaa ttgctgaact taaaaagcag tctaaacctg
541 tgaccacccc tgaagaaatt gcacagggtt ctacgatttc tgcaaacgga gacaaagaaa
601 ttggcaatat catctctgat gcaatgaaaa aagttggaag aaagggtgtc atcacagtaa
661 aggatggaaa aacactgaat gatgaattag aaattattga aggcattgaag tttgatcgag
721 gctatatctt tccatacttt attaatacat caaaagggtc gaaatgtgaa ttccaggatg
781 cctatgttct gttgagtga aagaaaattt ctagtatcca gtccattgta cctgctcttg
841 aaattgccaa tgctcaccgt aagccttttg tcataatcgc tgaagatgtt gatggagaag
901 ctctaagtac actcgtcttg aataggctaa aggttggtct tcaggttgtg gcagtcaagg
961 ctccagggtt tggtgacaat agaaagaacc agcttaaaga tatggctatt gctactggtg
1021 gtgcagtgtt tggagaagag ggattgacct tgaatcttga agacgttcag cctcatgact
1081 taggaaaagt tggagagggtc attgtgacca aagacgatgc catgctctta aaaggaaaag
1141 gtgacaaggc tcaaattgaa aaacgtattc aagaaatcat tgagcagtta gatgtcacia
1201 ctagtgaata tgaaaaggaa aaactgaatg aacggcttgc aaaaactttc gatggagtgg
1261 ctgtgctgaa ggttggtggg acaagtgatg ttgaagtga tgaaaagaaa gacagagtta
1321 cagatgccct taatgctaca agagctgctg ttgaagaagg cattgttttg ggaggggggt
1381 gtgccctcct tcgatgcatt ccagccttgg actcattgac tccagctaag gaagatcaaa
1441 aaattgggat agaaattatt aaaagaacac tcaaaattcc agcaatgacc attgctaaga
1501 atgcagggtg tgaaggatct ttgatagttg agaaaattat gcaaagttcc tcagaagttg
1561 gttatgatgc tatggctgga gattttgtga atatggtgga aaaaggaatc attgacccaa
1621 caaaggttgt gagaactgct ttattggatg ctgctgggtg ggccctctctg ttaactacag
1681 cagaagttgt agtcacagaa attcctaaag aagagaagga ccctggaatg ggtgcaatgg
1741 gtggaatggg aggtgggtatg ggaggtggca tgttctaact cctagactag tgctttacct
1801 ttattaatga actgtgacag gaagcccaag gcagtgttcc tcaccaataa cttcagagaa
1861 gtcagttgga gaaaatgaag aaaaaggctg gctgaaaatc actataacca tcagttactg
1921 gtttcagttg acaaaatata taatggttta ctgctgtcat tgtccatgcc tacagataat
1981 ttattttgta tttttgaata aaaaacattt gtacattcct gatactgggt acaagagcca
2041 tgtaccagtg tactgctttc aacttaaatc actgaggcat ttttactact attctgttaa
2101 aatcaggatt ttagtgcttg ccaccaccag atgagaagtt aagcagcctt tctgtggaga
2161 gtgagaataa ttgtgtacaa agtagagaag tatccaatta tgtgacaacc tttgtgtaat
2221 aaaaatttgt ttaaagttaa aaaaaaaaaa aaaaaaaa

```

FIGURE 7



ZPK (DNA sequence; SEQ ID NO:6)

```

1  agcatccgga gcgagctgc agcagcgccg ccttttgtgc tgcggccgcg gagccccga
61  gggcccagtg ttcacatca taccaggggc cagaggcgat ggcttgcctc catgagaccc
121 gaacaccctc tccttccttt gggggctttg tgtctaccct aagtgaggca tccatgcgca
181 agctggaccc agacacttct gactgcactc ccgagaagga cctgacgcct acccatgtcc
241 tgcagctaca tgagcaggat gcagggggcc cagggggagc agctgggtca cctgagagtc
301 gggcatccag agttcgagct gacgaggtgc gactgcagtg ccagagtggc agtggcttcc
361 ttgagggcct ctttggtgc ctgcgccttg tctggaccat gattggcaaa gcctactcca
421 ctgagcacia gcagcagcag gaagaccttt gggagggtccc ctttgaggaa atcctggacc
481 tgcagtgggt gggctcaggg gcccaggggtg ctgtcttctt ggggcgcttc cacggggagg
541 aggtggctgt gaagaagggt cgagacctca aagaaaccga catcaagcac ttgcgaaagc
601 tgaagcacc caacatcatc actttcaagg gtgtgtgcac ccaggctccc tgctactgca
661 tcctcatgga gttctgcgcc cagggccagc tgtatgaggt actgcgggct ggccgccttg
721 tcaccccttc cttactggtt gactgggtcca tgggcatcgc tgggtggcatg aactacctgc
781 acctgcacaa gattatccac agggatctca agtcacccaa catgctaata acctacgacg
841 atgtgggtga gatctcagat ttgggcaact ccaaggagct gagtgacaag agcaccaaga
901 tgtcctttgc agggacagta gcctggatgg cccttgaggt gatccgcaat gaacctgtgt
961 ctgagaaggt cgacatctgg tcctttggcg tgggtctatg ggaactgctg actggtgaga
1021 tcccctacaa agacgtagat tcctcagcca ttatctgggg tgtgggaagc aacagtctcc
1081 atctgcccgt gccctccagt tgcccagatg gtttcaagat cctgcttcgc cagtgtcggg
1141 atagcaaacc acgaaatcgc ccatcattcc gacagatcct gctgcatctg gacattgcct
1201 cagctgatgt actctccaca ccccaggaga cttactttaa gtcccaggca gagtggcggg
1261 aagaagtaaa actgcacttt gaaaagatta agtcagaagg gacctgtctg caccgcctag
1321 aagaggaact ggtgatgagg aggaggagg agctcagaca cgccctggac atcaggggagc
1381 actatgaaag gaagctggag agagccaaca acctgtatat ggaacttaat gccctcatgt
1441 tgcagctgga actcaaggag agggagctgc tcaggcgaga gcaagcttta gagcggagggt
1501 gcccaggcct gctgaagcca cacccttccc ggggcctcct gcatggaaac acaatggaga
1561 agcttatcaa gaagaggaat gtgccacaga agctgtcacc ccatagcaaa aggccagata
1621 tcctcaagac ggagtctttg ctccctaaac tagatgcagc cctgagtggg gtggggcttc
1681 ctgggtgtcc taaggccccc ccctcaccag gacggagtgc ccgtggcaag acccgtcacc
1741 gcaaggccag cgccaagggg agctgtgggg acctgcctgg gcttcgtaca gctgtgccac
1801 ctgatgacat tggaggacca ggaagcccaa ggggcctagg agggggaccc tcagcctggg
1861 aggcctgccc tcccgccttc cgtgggcttc atcatgacct cctgctccgc aaaatgtctt
1921 catcgtcccc agacctgctg tcagcagcac tagggtcccg gggccggggg gccacaggcg
1981 gagctgggga tcctggctca ccacctccgg cccggggtga caccaccaca agtgagggct
2041 cagcccctgg ctccaccagc ccagattcac ctgggggagc caaaggggaa ccacctcttc
2101 cagtaggggc tggtaagggt gtggggcttc tgggaactgg aagggaaggg acctcaggcc
2161 ggggaggaag ccgggtctgg tcccagcact tgacccagc tgcactgctg tacagggctg
2221 ccgtcaccgg aagtcagaaa cgtggcatct catcggaaga ggaggaagga gaggtagaca
2281 gtgaagtaga gctgacatca agccagaggt ggctcagag cctgaacatg cgccagtcac
2341 tatctacctt cagctcagag aatccatcag atggggagga aggcacagct agtgaacctt
2401 ccccatgtgg cacacctgaa gttggcagca ccaacactga tgagcgcca gatgagcgtt
2461 ctgatgacat gtgctcccag ggctcagaaa tcccactgga cccacctcct tcagaggtca
2521 tccctggccc tgaaccagc tcctgcccc ttccacacca ggaacttctc agagagcggg
2581 gccctcccaa ttctgaggac tcagactgtg acagactga attggacaac tccaacagcg
2641 ttgatgcctt ggggccccca gcttcccttc ctccatgaaa gccactcgta ttcttgtac
2701 atagagaaat atttatatgg attatatata tatacatata tatatatata tgcgccacat
2761 aatcaacaga aagatggggc tgtcccagcc gtaagtcagg ctgaggggag actgatcccc
2821 tgaccaattc acctgataaa ctctagggac actggcagct gtggaaatga atgaggcaca
2881 gccgtagagc tgtggctaag ggcaagcccc ttctgcccc accccattcc ttatattcag
2941 caagcaacaa ggcaatagaa aagccagggt tgtctttata ttctttatcc ccaaataata
3001 gggggtgggg ggagggggcg tgggaggggc aggagagaaa accacttaga ctgcactttt
3061 ctgttccgtt tactctgttt acacattttg cacttgggag gagggaggct aaggctgggt
3121 cctcccctct gaggtttctc aggtggcaat gtaactcatt tttttgtccc accatttatc

```

FIGURE 8

3181 ttctctgccc aagccctgtc ttaaggccca gggggagggt aggagactga tagcatgtga  
3241 tggctcaggc tgaagaaccg ggggtgctgt taagtccctg cttttatcct ggtgcctgat  
3301 tggggtgggg actgtcctac tgtaaccctt gtgaaaaacc ttgaaatata acactccatg  
3361 cagga

FIGURE 8 (cont.)

SCD (protein; SEQ ID NO:7)

MPAHLQDDISSSYTTTTTITAPPSRVLQNGGDKLETMPLYLEDDIRPDIKDDIYDPTY  
KDKEGSPSPKVEYVWRNIILMSLLHLGALYGITLIPTCKFYTWLWGVFYYFVSALGITAG  
AHRLWSHRSYKARLPLRLFLIIANTMAFQNDVYEWARDHRAHHKFSETHADPHNSRRGF  
FFSHVGWLLVRKHPAVKEKGSTLDLSDLEAEKLVMFQRRYYKPGLLLMCFILPTLVWPY  
FWGETFQNSVFVATFLRYAVVLNATWLVNSAAHLFGYRPHYDKNISPRENILVSLGAVGE  
GFHNYHHSFPYDYSASEYRWHINFTTFFIDCMAALGLAYDRKKVSKAAILARIKRTGDG  
NYKSG

FIGURE 9

CA12 (Protein; SEQ ID NO:8)

MRRSLHAAVLLLVLKEQPSSPAPVNGSKWTYFGPDGENSWSKYPSCGGLLQSPID  
LHSDILQYDASLTPLFQGYNLSANKQFLLTNNGHSVKLNLPSDMHIQGLQSRYSATQL  
HLHWGNPNDPHGSEHTVSGQHFAAELHIVHYNSDLYPDASTASNKSEGLAVLAVLIEMG  
SFNPSYDKIFSHLQHVKYKGQEAFFVPGFNIEELLPERTAEYYRYRGSLLTPPCNPTVLW  
TVFRNPVQISQEQLLALETALYCTHMDPSPREMINNFRQVQKFDERLVYTSFSQVQVC  
TAAGLSLGIILSLALAGILGICIVVVSIWLFRRKSIKKGDNKGVIYKPATKMETEAHA

FIGURE 10

PIK3R4 (Protein; SEQ ID NO:9)

MGNQLAGIAPSQILSVESYFSDIHDFEYDKSLGSTRFFKVARAKHREGLVVVKVFAIQD  
PTLPLTSYKQELEELKIRLNSAQNCLEPFQKASEKASEKAAMLFRQYVRDNLVDRISTRP  
FLNNIEKRWIAFQILTAVDQAHKSGVRHGDIKTENVMVTSWNWVLLTDFASFKPTYLPE  
DNPADFNYFFDTSRRRTCIIAPERFVDGGMFATELEYMRDPSTPLVDLNSNQRTRGELK  
RAMDIFSAGCVIAELFTEGVPLFDLSQLLAYRNGHFFPEQVLNKIEDHSIRELVTQMIH  
REPDKRLEAEDYLNKQQRGNAPFEIFYTFLOPYMAQFAKETFLSADERILVIRKDLGNI  
HNLCGHDLPEKAEGEPKENGLVILVSVITSCLQTLKYCDKLALELILHLAPRLSVEI  
LLDRITPYLLHFSNDSVPRVRAEALRTLTKVLALVKEVPRNDINIYPEYILPGIAHLAQ  
DDATIVRLAYAENIALLAETALRFLLELVQLKNLNMENDPNNEEIDEVTHPNGNYDTELQ  
ALHEMVQQKVVTLLSDPENIVKQTLMEGITRLCVFFGRQKANDVLLSHMITFLNDKND  
WHLRGAFFDSIVGVAAYVGWQSSSILKPLLQOGLSDAEFVIVKALYALTCMCQLGLLQ  
KPHVYEFASDIAPFLCHPNLWIRYGAVGFITVVARQISTADVYCKLMPYLDPIYTQPII  
QIERKLVLLSVLKEPVSRISFDYALRSKDITSLFRHLHMRQKKRNGSLPDCPPPEDPAI  
AQLLKKLLSQGMTEEEEDKLLALKDFMMKSNKAKANIVDQSHLHDSQKGVIDLALGI  
TGRQVDLVKTKQEPDDKRARKHVKQDSNVNEEWKSMFGSLDPPNMPQALPKGSDQEVII  
TGKPPRSESSAGICVPLSTSSQVPEVTTVQNKKPVIPVLSSTILPSTYQIRITTCKTEL  
QQLIQKREQCNAERIAKQMMENAWESEKPPPPGWRPKGLLV AHLHEHKS AVNRIRVSD  
EHS LFATCSNDGT VKI WNSQKMEGKTTTTRSILTYSRIGGRVKTLTFCQGSHYLAIASD  
NGAVQLLGIEASKLPKSPKIHPLQSRILDQKEDGCVDMHMFNSGAQSVLAYATVNGSL  
VGWDLRSSSNWTLKHDLKSGLITSFAVDIHQCWLCIGTSSGTMACWDMRFQLPISSHC  
HPSRARIRRLSMHPLYQSWVIAAVQGNNEVSMWDMETGDRRFTLWASSAPPLSELQPS  
HSVHGIYCSPADGNPILLTAGSDMKIRFWDLAYPERSYVAGSTSSPSVSYYRKIEGT  
EVVQEIQNKQKVGPSDDTPRRGPESLPVGHHDII TDVATFQTQGFIVTASRDGIVKVK  
K

FIGURE 11

PLD3 (Protein; SEQ ID NO:10)

MTQLFLWEYGDLHLFGPNQRPAPCYDPCEAVLVESIPEGLDFPNASTGNPSTSQAWLGL  
LAGAHSSLDIASFYWTLTNNDTHTQEPSAQQGEEVLRQLQTLAPKGVNVRIAVSKPSGP  
QPQADLQALLQSGAQVRMVDQMQLTHGVLHTKFWVVDQTHFYLG SANMDWRS LTQVKEL  
GVVMYNCSCLARDLTKIFEAYWFLGQAGSSIPSTWPRFYDTRYNQETPMEICLNGTPAL  
AYLASAPPPLCPSGRTPDLKALLNVVDNARSFIYVAVMNYLPTLEFSHPHRFWPAIDDG  
LRRATYERGVKVRLLISCWGHSEPSMRAFLLSLAALRDNHTHSDIQVKLFVVPAD E AQA  
RIPYARVNH NKYMVTERATYIGTSNWSGNYFTETAGTSLLV TQNGRGGLRSQLEAIFLR  
DWDSPIYHDLDT SADS VGNACRLL

HSPD1 (Protein; SEQ ID NO:11)

MLRLPTVFRQMRPVSRLAPHLTRAYAKDVKFGADARALMLQGVDLLADAVAVTMGPKG  
RTVIEQSWGSPKVT KDGVTVAKSIDLKDKYKNIGAKLVQDVANNTNEEAGDGT TTTATV  
LARSIAKEGF EKISKGANPVEIRRGVMLAVDAVIAELKKQSKPVTTPEEIAQVATISAN  
GDKEIGNIISDAMKKVGRKGVITVKDGKTLNDELEIIEGMKFDRGYISPYFINTSKGQK  
CEFQDAYVLLSEKKISSIQSIVPALEIANAHRKPLVIIAEDVDGEALSTLVNRLKVGL  
QVVAVKAPGFGDNRKNQLKDMAIATGGAVFGEEGLTLNLEDVQPHDLGKVGEVIVTKDD  
AMLLKGKGDKAQIEKRIQEIEQLDVT TSEYEKEKLNERLAKLSDGVAVLKVGGTSDVE  
VNEKKDRVTDALNATRAAVEEGIVLGGGCALLRCIPALDSLTPANEDQKIGIEIIKRTL  
KIPAMTIAKNAGVEGSLIVEKIMQSSSEVGYDAMAGDFVNMVEKGIIDPTKVVRTALLD  
AAGVASLLTTAEVVVTEIPKEEKDPGMGAMGGMGGMGGGMF

FIGURE 13

ZPK (Protein; SEQ ID NO:12)

MACLHETRTSPSPSFGGFVSTLSEASMRKLDPDTS DCTPEKDLTPTHVLQLHEQDAGGPG  
GAAGSPESRASRVRADEVRLQCQSGSGFLEGLFGCLRPVWTMIGKAYSTEHKQQQEDLW  
EVPFEEILDQLQWVGSGAQGAVFLGRFHGEEVAVKKVRDLKETDIKHLRKLKHPNIIITFK  
GVCTQAPCYCILMEFCAQQQLYEVLRAGRPVTPSLLVDWSMGIAGGMNYLHLHKIIHRD  
LKSPNMLITYDDVVKISDFGTSKELSDKSTKMSFAGTVAWMAPEVIRNEPVSEKVDIWS  
FGVVLWELLTGEIPYKDVDSSAIIWGVGSNSLHLPVPSSCPDGFKILLRQCWNSKPRNR  
PSFRQILLHLDIASADVLSTPQETYFKSQAEWREEVKLHFEEKIKSEGTCLHRLEEELVM  
RRREELRHALDIREHYERKLERANNLYMELNALMLQLELKERELLRREQALERRCPGLL  
KPHPSRGLLHGNTMEKLIKRNVPQKLSPHSKRPDILKTESLLPKLDAALSGVGLPGCP  
KAPPSPGRSRRGKTRHRKASAKGSCGDLPLRTAVPPHEPGGPGSPGGLGGGPSAWEAC  
PPALRGLHHDLLLRLKMSSSSPDLLSAALGSRGRGATGGAGDPGSPPPARGDTPPSEGSA  
PGSTSPDSPGGAKEPPPPVGPGEVGLLGTGREGTSGRGGSRAGSQHLTPAALLYRAA  
VTRSQKRGISSEEEEGEVDSEVELTSSQRWPQSLNMRQSLSTFSSENPSDGEETASEP  
SPSGTPEVGSTNTDERPDERSDDMCSQGSEIPLDPPPSEVI PGPEPSSLPIPHQELLRE  
RGPPNSEDSDCDSTELDNSNSVDALRPPASLPP

FIGURE 14



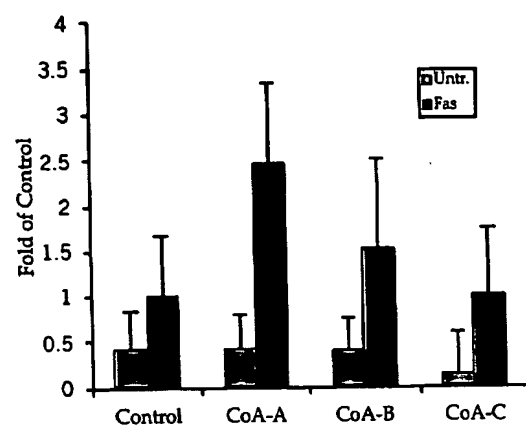


FIGURE 15

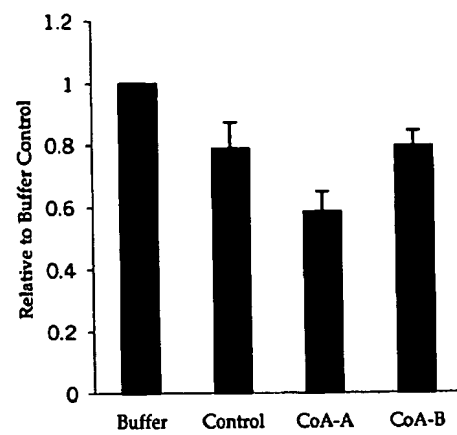


FIGURE 16

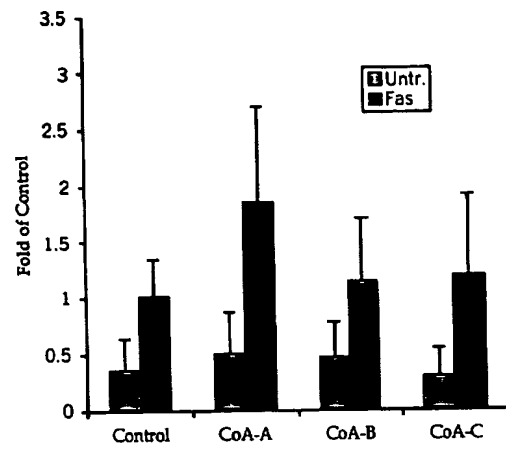


FIGURE 17

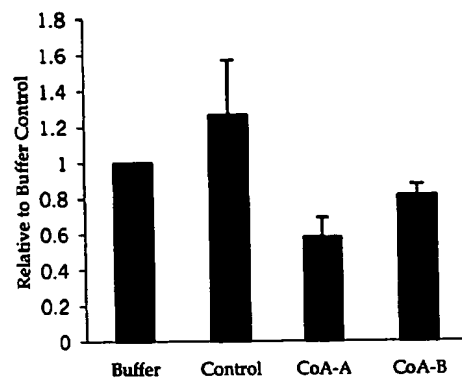


FIGURE 18

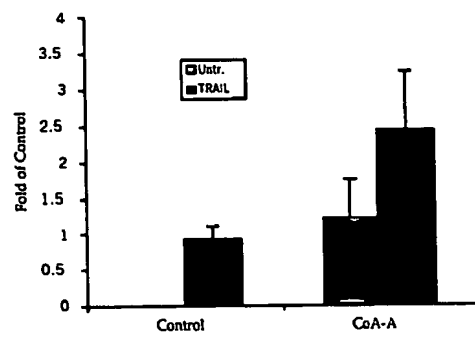


FIGURE 19

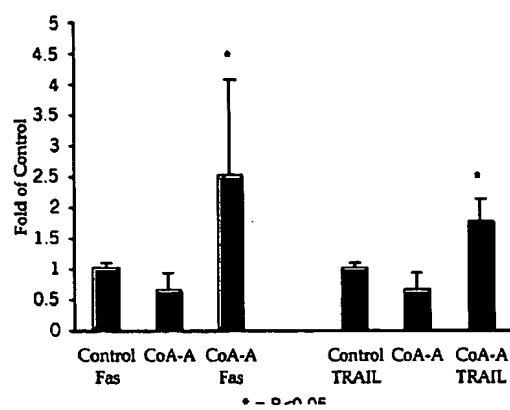


FIGURE 20

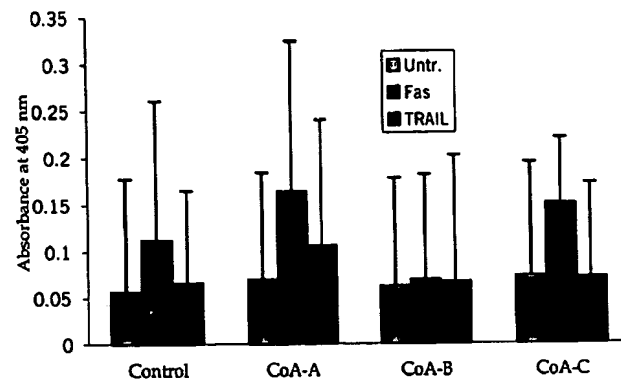


FIGURE 21

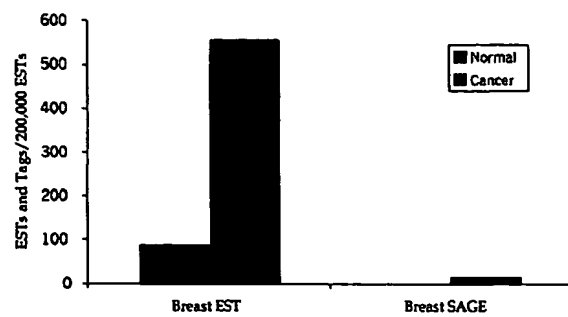


FIGURE 22



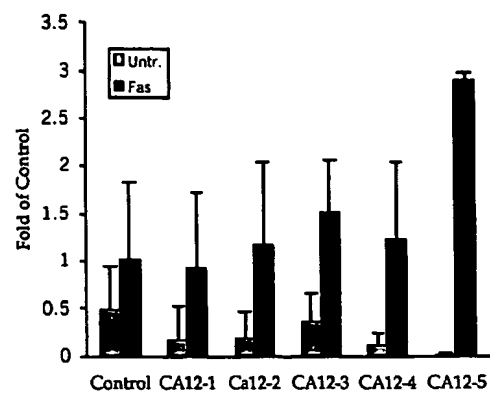


FIGURE 23

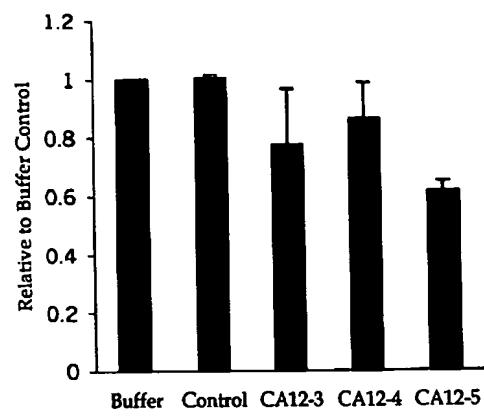


FIGURE 24

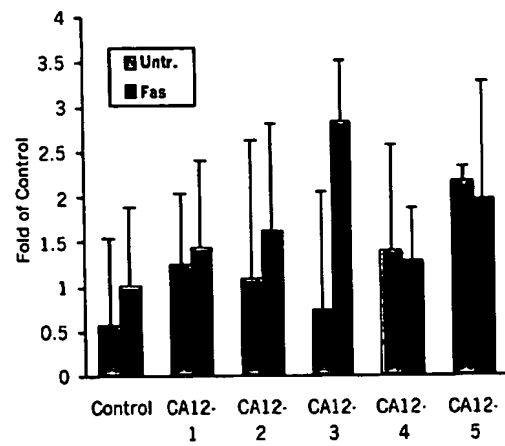


FIGURE 25

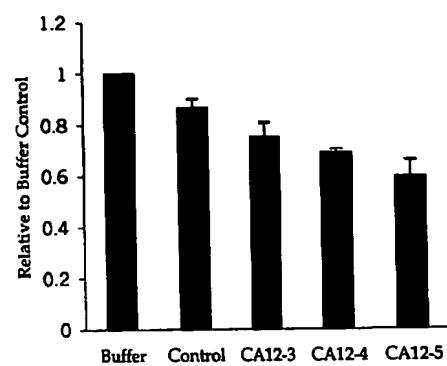


FIGURE 26

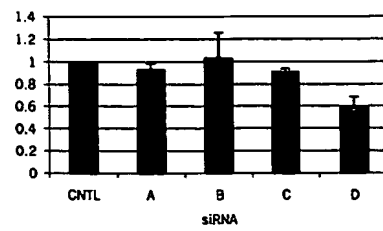


FIGURE 27

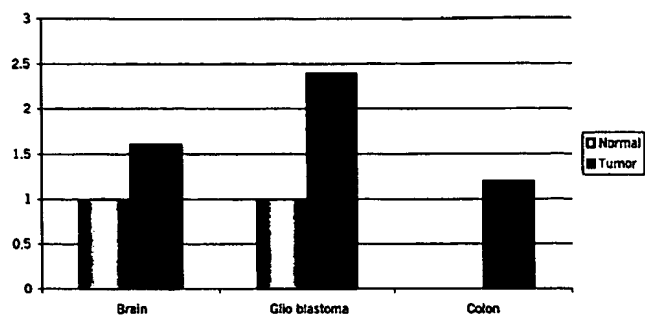


FIGURE 28

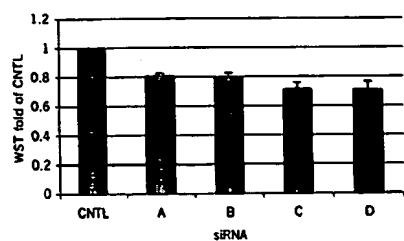


FIGURE 29A

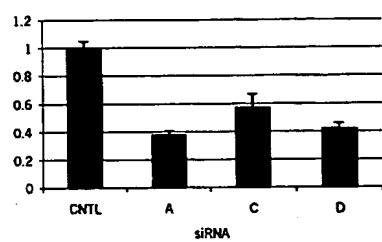


FIGURE 29B

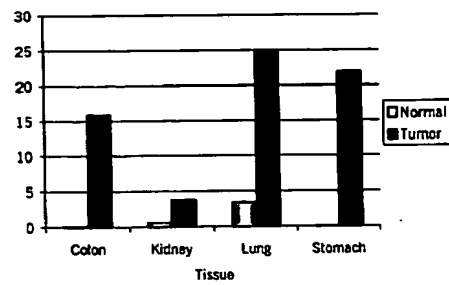


FIGURE 30



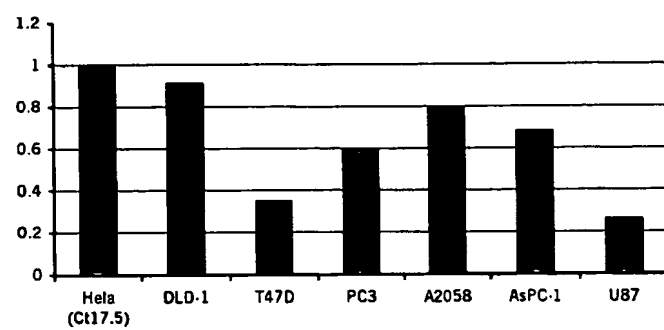


FIGURE 31

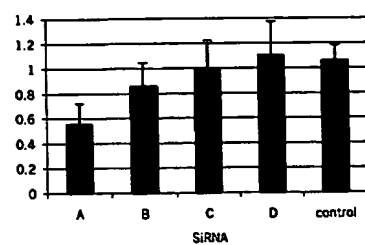


FIGURE 32A

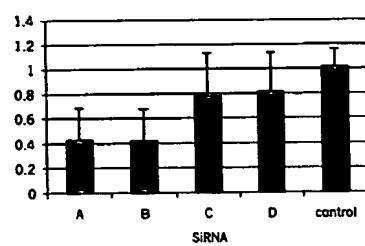


FIGURE 32B

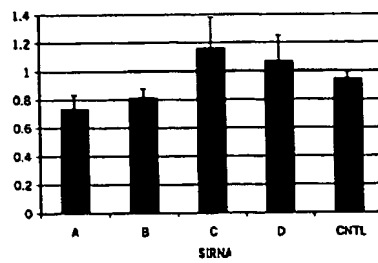


FIGURE 33A

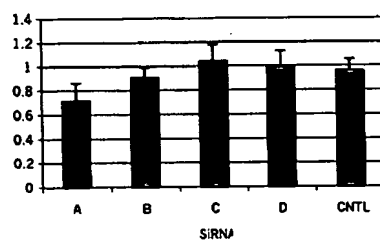


FIGURE 33B

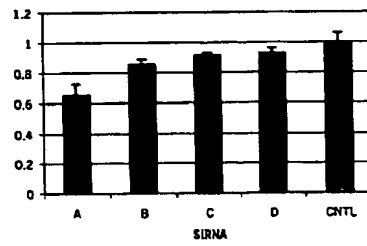


FIGURE 34

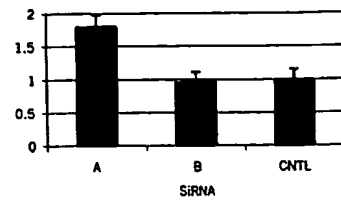


FIGURE 35

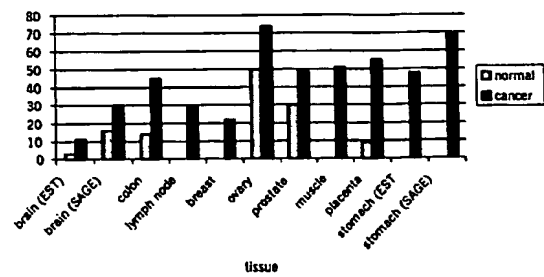


FIGURE 36

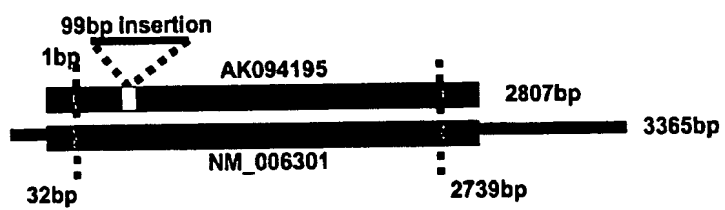


FIGURE 37

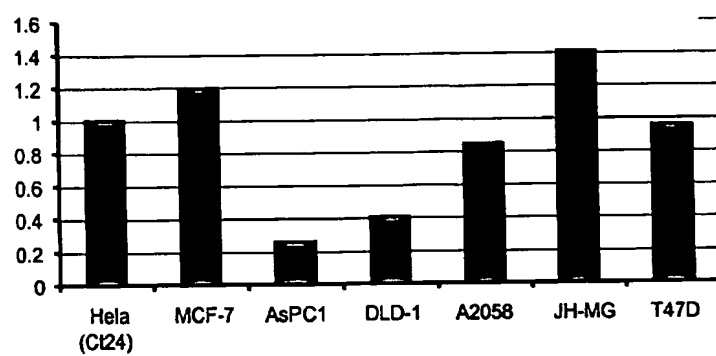


FIGURE 38



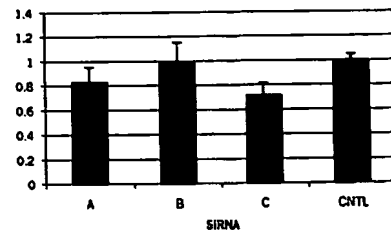


FIGURE 39A

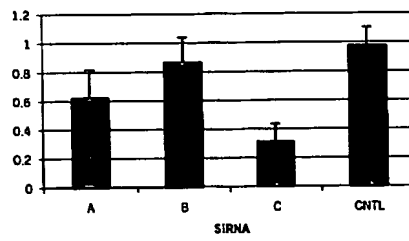


FIGURE 39B

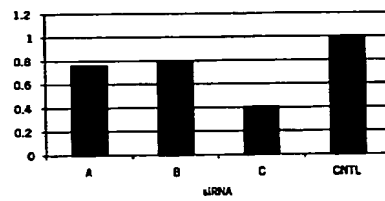


FIGURE 39C

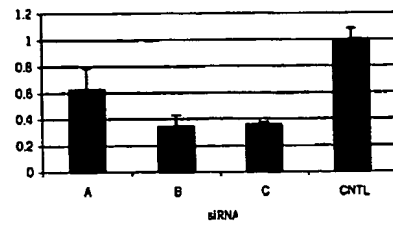


FIGURE 40A

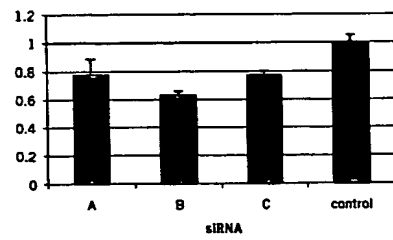


FIGURE 40B

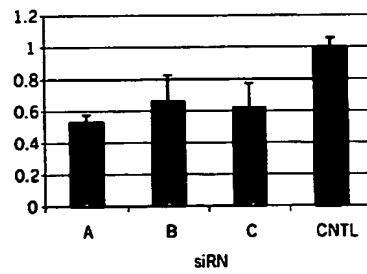


FIGURE 40C

# DNA Sequence of Variant 1 of ZPK (SEQ ID NO:40)

```

1 cttttgtgct gcggccgagg agcccccgag ggcccagtgt tcaccatcat accagggggc
61 agaggcgatg gcttgccctc atgagacccg aacacctctt ccttcctttg ggggctttgt
121 gtctacccta agtgaggcat ccatgcgcaa gctggaccca gacacttctg actgcactcc
181 cgagaaggac ctgacgccta cccagtgtgt acttcgagat gtggtacccc ttggtgggca
241 ggggtggggg gggcccagcc cctccccagg tggagagccg cccctgagc cctttgccaa
301 cagtgtcctg cagctacatg agcaggatgc agggggccca gggggagcag ctgggtcacc
361 tgagagtcgg gcatccagag ttcgagctga cgaggtgcga ctgcagtgcc agagtggcag
421 tggtttcctt gagggcctct ttggctgcct gcgccctgtc tggaccatga ttggcaaaagc
481 ctactccact gagcacaagc agcagcagga agacctttgg gaggtcccct ttgaggaaat
541 cctggacctg cagtgggtgg gctcaggggc ccagggtgct gtcttcctgg ggcgcttcca
601 cggggaggag gtggctgtga agaagggtgc agacctcaa gaaaccgaca tcaagcactt
661 gcgaaagctg aagcacccca acatcatcac tttcaagggt gtgtgcaccc aggtccctg
721 ctactgcata ctcatggagt tctgcgccc gggccagctg tatgaggtac tgcgggctgg
781 ccgccctgtc acccctcctt tactgggtga ctggtccatg ggcacgctg gtggcatgaa
841 ctacctgcac ctgcacaaga ttatccacag ggatctcaag tcaccaaca tgctaatac
901 ctacgacgat gtggtgaaga tctcagattt tggcacttcc aaggagctga gtgacaagag
961 caccaagatg tcctttgcag ggacagtagc ctggatggcc cctgaggtga tccgcaatga
1021 acctgtgtct gagaaggctg acatctggtc ctttggcgtg gtgctatggg aactgctgac
1081 tggtagatc cctacaaaag acgtagattc ctcagccatt atctggggtg tgggaagcaa
1141 cagtctccat ctgccgtgc cctccagttg cccagatggt tcaagatcc tgcttcgcca
1201 gtgctggaat agcaaacac gaaatcgctc atcattccga cagatcctgc tgcactctga
1261 cattgcctca gctgatgtac tctccacacc ccaggagact tactttaagt cccaggcaga
1321 gtggcgggaa gaagtaaaac tgcactttga aaagattaag tcagaaggga cctgtctgca
1381 ccgcctagaa gaggaactgg tgatgaggag gaggaggag ctcagacacg ccctggacat
1441 caggagcac tatgaaagga agctggagag agccaacaac ctgtatatgg aacttaatgc
1501 cctcatgttg cagctggaac tcaaggagag ggagctgctc aggcgagagc aagctttaga
1561 gcggaggtgc ccaggcctgc tgaagccaca cccttcccgg ggctcctgc atggaaaac
1621 aatggagaag cttatcaaga agaggaatgt gccacagaag ctgtcacccc atagcaaaag
1681 gccagatata ctcaagacgg agtctttgct ccctaaacta gatgcagccc tgagtggggt
1741 ggggcttctt ggggtgtcta agggccccc ctcaccagga cggagtcgcc gtggcaagac
1801 ccgtcacccg aaggccagcg ccaaggggag ctgtggggac ctgcctgggc ttcgtacagc
1861 tgtgccacc catgaacctg gaggaccagg aagcccaggg ggcctaggag ggggaccctc
1921 agcctgggag gcctgcccct ccgccctccg tgggcttcat catgacctcc tgctccgcaa
1981 aatgtcttca tcgtccccag acctgctgtc agcagcacta ggggtcccgg gcccgggggc
2041 cacaggcgga gctggggatc ctggctcacc acctccggcc cggggtgaca cccaccaag
2101 tgagggtcga gccctggct ccaccagccc agattcacct gggggagcca aagggaacc
2161 acctcctcca gtagggcctg gtgaagggtg ggggcttctg ggaactggaa ggggaaggac
2221 ctcaggcccg ggaggaagcc gggctgggtc ccagcacttg accccagctg cactgctgta
2281 cagggtgcc gtcacccgaa gtcagaaacg tggcatctca tcggaagagg aggaaggaga
2341 ggtagacagt gaagtagagc tgacatcaag ccagaggtgg cctcagagcc tgaacatgcg
2401 ccagtcacta tctaccttca gctcagagaa tccatcagat ggggaggaag gcacagctag
2461 tgaaccttcc cccagtggca cacctgaagt tggcagcacc aacactgatg agcgccaga
2521 tgagcgggtc gatgacatgt gctcccaggg ctcagaaatc ccactggacc cacctcctc
2581 agaggtcact cctggccctg aaccacgctc cctgcccatt ccacaccagg aacttctcag
2641 agagcggggc cctcccaatt ctgaggactc agactgtgac agcactgaat tggacaactc
2701 caacagcgtt gatgccttgc ggccccagc ttccctccct ccatgaaagc cactcgtatt
2761 ccttgtagat agagaaatat ttatataaat tatatatata tacatat

```

FIGURE 41

**Amino Acid Sequence of Variant 1 of ZPK (SEQ ID NO:41)**

MACLHETRTSPSPSGGFVSTLSEASMRKLDPDTSDCTPEKDLTPTQCVLRDVVPLGGQGGGGPSPSPGGE  
PPPEPFANSVLQLHEQDAGGPGGAAGSPESRASRVRADEVRLQCQSGSGFLEGLFGCLRPVWTMIGKAYS  
TEHKQQQEDLWEVPFEEILDQWVGSGAQGAVFLGRFHGEEVAVKKVRDLKETDIKHLRKLKHPNIIITFK  
GVCTQAPCYCILMEFCAQGQLYEVLRAGRPVTPSLLVDWSMGIAGGMNYLHLHKI IHRDLKSPNMLITYD  
DVVKISDFGTSKELSDKSTKMSTKMSFAGTVAWMAPEVIRNEPVSEKVDIWSFGVVLWELLTGEI PYKDVDSSA  
IIWGVGSNSLHLPVPSSCPDGFKILLRQCWNSKPRNRSSFRQILLHLDIASADVLSTPQETYFKSQAWEW  
EEVKLHFEEKIKSEGTCLHRLEEELVMRRREELRHALDIREHYERKLERANNLYMELNALMLQLELKEREL  
LRREQALERRCPGLLKPHPSRGLLHGNTMEKLIKRNVPQKLSPHSKRPDILKTESLLPKLDAALSGVGL  
PGCPKGGPPSPGRSRRGKTRHRKASAKGSCGDLPLRTAVPPHEPGGPGSPGGLGGGPSAWEACPPALRGL  
HHDLLLRLKMSSSSPDLLSAALGSRGRGATGGAGDPGSPPPARGDTPPSEGSAPGSTSPDSPGGAKGEPPP  
PVGPGEGVGLLGTGREGTSGRGGSRAGSQHLTPAALLYRAAVTRSQKRGISSEEEEGEVDSEVELTSSQR  
WPQSLNMRQSLSTFSSENPSDGEEGTASEPSPSGTPEVGSTNTDERPDERSDDMCSQGSEIPLDPPPSEV  
IPGPEPSSLPIPHQELLRERGGPPNSEDSDCDSTELDNSNSVDALRPPASLPP

**FIGURE 42**

## PROVISIONAL PATENT APPLICATION TRANSMITTAL

Docket No.:

P-IMM 1008

COMMISSIONER FOR PATENTS  
Mail Stop: Provisional Application  
P. O. Box 1450  
Alexandria, VA 22313-1450

22389 U.S. PTO

60/500281



CERTIFICATE OF MAILING BY "EXPRESS MAIL"

"EXPRESS MAIL" MAILING LABEL NUMBER: EP336774120US

DATE OF DEPOSIT: September 4, 2003

I HEREBY CERTIFY THAT THIS PAPER IS BEING  
DEPOSITED WITH THE UNITED STATES POSTAL SERVICE  
"EXPRESS MAIL POST OFFICE TO ADDRESSEE" SERVICE UNDER  
37 C.F.R. 1.10 ON THE DATE INDICATED ABOVE, AND IS  
ADDRESSED TO: COMMISSIONER FOR PATENTS, MAIL STOP: PROVISIONAL  
APPLICATION, P.O. BOX 1450, ALEXANDRIA, VA 22313-1450.

David I. Spolter  
(TYPED OR PRINTED NAME OR PERSON MAILING PAPER OR FEE)

  
(SIGNATURE OF PERSON MAILING PAPER OR FEE)

This is a request for filing a provisional patent application  
under 37 CFR 1.53(c).

Title of the Invention: METHODS OF IDENTIFYING AGENTS THAT  
INHIBIT THE GROWTH OF CANCER CELLS

Inventors:

Given Name	Surname	Residence
Xiuyuan	Hu	San Diego, California, U.S.
Henry	Li	Carlsbad, California, U.S.
Ning	Ke	San Diego, California, U.S.
Mirta	Grifman	San Diego, California, U.S.
Cheryl	Rogers	San Diego, California, U.S.
Kristin	Defife	San Diego, California, U.S.
Cellia	Habita	San Diego, California, U.S.
Wufang	Fan	San Diego, California, U.S.
Kristina	Rhoades	San Diego, California, U.S.
Philip	Tan	San Diego, California, U.S.
Richard	Tritz	San Diego, California, U.S.
Flossie	Wong-Staal	San Diego, California, U.S.

Provisional Application Transmittal  
Hu et al.  
Page 2

Enclosed are:

<u>X</u>		Return receipt postcard
<u>X</u>	<u>1</u>	Page application cover sheet
<u>X</u>	<u>101</u>	Pages of specification (includes claims and abstract)
<u>X</u>	<u>45</u>	Sheets of drawings

X A check in the amount of \$160.00 to cover the filing fee is enclosed.

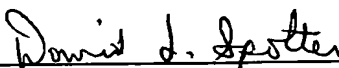
X The Commissioner is hereby authorized to charge fees under 37 CFR 1.16 and 1.17 which may be required or credit any overpayment to Deposit Account No. 501559. A duplicate copy of this sheet is enclosed.

Please address all future communications to:

David Spolter  
LAW OFFICE OF DAVID SPOLTER  
1590 Coast Walk  
La Jolla, California 92037  
Telephone: (858) 459-2934  
Facsimile: (858) 459-0698

Respectfully submitted,

Date: September 4, 2003

  
\_\_\_\_\_  
David I. Spolter  
Registration No. 36,933  
LAW OFFICE OF DAVID SPOLTER  
1590 Coast Walk  
La Jolla, California 92037

# Document made available under the Patent Cooperation Treaty (PCT)

International application number: PCT/US04/027672

International filing date: 24 August 2004 (24.08.2004)

Document type: Certified copy of priority document

Document details: Country/Office: US  
Number: 60/500,281  
Filing date: 04 September 2003 (04.09.2003)

Date of receipt at the International Bureau: 15 December 2004 (15.12.2004)

Remark: Priority document submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b)



World Intellectual Property Organization (WIPO) - Geneva, Switzerland  
Organisation Mondiale de la Propriété Intellectuelle (OMPI) - Genève, Suisse

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☒ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☒ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☒ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**